

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```
Run on:      Fri Sep 3 13:11:46 1999;  MasPar time 18.55 Seconds
```

Title: >US-09-120-044-3-COPY

```

Description: (1-4/1) from devil20044.pdf
Perfect Score: 3234

```

Sequence:

Scoring table:

Searched: 170751 seqs, 21266608 residues

```
Post-processing: Minimum Match 0%
                  Listing first 45 summaries
```

Database:

a-*genesec35*
1:part1 2:part2 3:part3 4:part4 5:parts 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Statistics:

Mean 35.314; Variance 169.601; scale 0.208

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	3232	99.9	471	1	R05923	Immunogenic pneumolys	6.18e-27
2	3056	94.5	471	1	R05924	Immunogenic pneumolys	3.57e-25
3	1416	43.8	480	6	R33842	Streptolysin O variant	5.45e-11
4	1398	43.2	480	6	R33841	Soluble streptolysin	2.13e-10
5	1391	43.0	571	2	R10376	Streptolysin O deriva	8.88e-10
6	1097	33.9	371	1	R06000	Bacteriophage lambda	7.61e-83
7	1162	11.3	113	3	R60952	Streptococcus pneumoniae	7.90e-20
8	112	3.5	11	17	R08135	Listeriolysin O haemo	5.96e-00
9	108	3.3	299	33	W60976	Streptococcus pneumoniae	1.10e-01
10	107	3.3	1228	25	W22862	Bacillus stearothermo	1.28e-01
11	107	3.3	1228	17	R77673	S-layer protein encod	1.28e-01
12	101	3.1	148	29	W55347	H. pylori ORF 01c116	3.16e-01
13	101	3.1	188	29	W55348	H. pylori ORF 28g9a303	3.16e-01
14	101	3.1	667	22	W20753	H. pylori transporter	3.16e+01
15	96	3.0	458	8	R41781	Rabbit HbI protein wh	6.16e+01
16	94	2.9	114	2	R07691	Myobacterial 65kD ant	8.85e+01

17	93	2.9	459	20	W04722	Aromatic acyl transferase	1.02e+07
18	94	2.9	555	33	W30676	Glycerol dehydratase	8.85e+01
19	94	2.9	555	32	W60264	Klebsiella pneumoniae	8.85e+01
20	95	2.9	649	38	W81977	Escherichia sp. B3 prot	7.65e+01
21	95	2.9	906	35	W17290	Potato starch branchi	7.65e+01
22	95	2.9	906	35	W69300	HSV-2 strain S85 cont	8.85e+01
23	94	2.9	1057	36	W72067	HSV-2 strain S85 cont	8.85e+01
24	94	2.9	1114	36	W72205	HSV-2 strain S85 cont	8.85e+01
25	94	2.9	1114	36	W72096	HSV-2 strain S85 cont	8.85e+01
26	93	2.9	1248	24	W18783	Human multiple regula	1.02e+02
27	90	2.8	210	15	R80623	Coho salmon growth ho	1.57e+02
28	91	2.8	226	20	W11582	Streptomyces pristina	1.36e+02
29	89	2.8	355	12	R64222	Human 38kDa FK-506 bi	1.81e+02
30	90	2.8	493	3	R13992	P. falciparum sporozo	1.57e+02
31	92	2.8	527	2	R05079	Human monomamine oxida	1.18e+02
32	92	2.8	527	33	W61278	Monamine oxidase A.	1.18e+02
33	90	2.8	556	17	R96737	A. niger Bo-1 carboxy	1.18e+02
34	92	2.8	625	23	W18008	Arabidopsis CER1 prot	1.57e+02
35	89	2.8	695	23	W19490	APP695 mutant A-beta	1.57e+02
36	90	2.8	713	3	R13991	Plasmodium falciparum	1.57e+02
37	89	2.8	751	23	W19505	APP751 mutant A-beta	1.81e+02
38	89	2.8	751	23	W19492	APP770 mutant A-beta	1.81e+02
39	89	2.8	770	23	W19506	TATA box binding prot	1.36e+02
40	89	2.8	1066	17	R91300	TFE.	1.36e+02
41	92	2.8	1093	3	R42818	Human multiple regula	1.18e+02
42	92	2.8	1444	24	W19785	Human multiple regula	1.18e+02
43	92	2.8	1726	24	W19784	P. falciparum liver s	1.18e+02
44	89	2.8	1786	23	W24790	Chicken cytochrome b	1.36e+02
45	91	2.8	1810	16	R94563		

ALIGNMENTS

ID	RESULT	1
AC	R05923 standard; protein; 471 AA.	
DT	R05923:	
DE	23-NOV-1990 (first entry)	
KW	Immunogenic pneumolysin variant.	
OS	Pneumolysin; vaccine; pneumonia; meningitis; bacteraemia; ds.	
PN	Streptococcus pneumoniae.	
PD	MO906951-A.	
PF	28-JUN-1990.	
PR	15-DEC-1989; AU0539.	
PA	16-DEC-1988; AU-001989.	
PI	(PATO)/ PATON J C.	
PI	PATON JC, HANSMAN DJ, BOUTNOIS GJ, ANDREW PW, MITCHELL TJ,	
DR	WALKER JB:	
DR	WPI: 90-224494/29.	
PT	N-PSDB: Q05270.	
PT	New non-toxic, immunogenic mutants of pneumolysin - useful in	
PT	protective vaccines against Streptococcus pneumoniae, and DNA	
PS	sequences encoding them.	
CC	Claim 5; Page 11; 25pp; English.	
CC	Vaccines are non-toxic and antigenic to wild type pneumolysin,	
CC	making them useful in vaccination against pneumonia and associated	
CC	meningitis, bacteraemia etc.	
SQ	Sequence 471 AA:	
	Query Match	99.9%; Score 3232; DB 1; Length 471;
	Best Local Similarity	97.9%; Pred. No. 6.18e-273;
	Matches	461; Conservative 1; Mismatches 9; Indels 0; Gaps
Db	1	markandfilimnydkkkllthggesienrfikbegnqldpdefvylærkkslstntedi 60
QY	1	MAKKAANDFILAMANDKRLKLLTHGGESENRFKKEGNQLPDEFVYXERKRKSLSTNTSDI 60
Db	61	svtbatndslrlypgallvydellennpdlavtraemtysidlpglassdsfiqvedpsn 120
QY	61	xvtatntdsrlrgallvydellxennpftlladvrapmtysidlpglassdsfiqvedpsn 120
Db	121	ssvrgavndllakwbdyqgvnvpamryekltahsmeglkvktfgsgfektgnsldidf 180
QY	121	SSVRGA VNDLLAKWHQDYQGVNNVPARXQYERKLTASHMSQLVKFGSDFEKTGNSLDIDF 180

Db 181 nswhsgekgqivnfgkqlyyrvsvdavnpgdfqdvteedlkqrgiseprrlyisv 240
 |||||
 Qy 181 NSVHSEKQIOIVNKKOITVSDAVKKNPGDFQDVTVEEDLKORGISAEPLVYISXV 240
 Db 241 aygrgyvllkettssdeveaafaelikgvkvaoptekqjldntevkavllggdpsga 300
 |||||
 Qy 241 AYGRGYVLLKETTSSXSEVEAFAELIKGVKVAOPTEKQJLDNTEVKAVLLGGDPSSGA 300
 Db 301 rvtgkvdmvedlqgesrftadhpglpisytstfirdnvaatfgnstdyvetkayrn 360
 |||||
 Qy 301 RVTGKVDMEVDLQGESRFTADHPGLPISTYTSFLRDNVATFQNSTDYETVKAYRN 360
 Db 361 gdlldhsagayvaayitfdelsydhgkveitpkawdrngqdltahtstipkgnvrn 420
 |||||
 Qy 361 GDLLDHSGAYVAAYITWNLSDYHOGKEVITPKAMDNGDULTAHTSTIPKGNVRN 420
 Db 421 lsvkirectglawemwrtvyecktdplvrkrtstsgwtltpyvedkveend 471
 |||||
 Qy 421 LSVKIRECTGLAWEMWRTYETKTDPLVRKRRTISMGTTLPYQVEDKVEEND 471

RESULT 2

ID R05924 standard; protein: 471 AA.
 AC R05924:
 DE 23-NOV-1990 (first entry)
 DE Immunogenic pneumolysin variant.
 KW Pneumolysin; vaccine; pneumonia; meningitis; bacteriaemia; ds.
 OS Streptococcus pneumoniae.
 FH Key location/Qualifiers
 FT misc_difference 428
 FT /label=G, A, S
 FT misc_difference 435
 FT /label=Q, D
 PN WO9006951-A.
 PD 28-JUN-1990.
 PE 15-DEC-1989; AU0539.
 PR 16-DEC-1989; AU-001989.
 PA (PATON/J. PATON J. C.
 PI PATON JC, HANSMAN DJ, BOULNOIS GJ, ANDREW PW, MITCHELL TJ,
 PI WALKER JA.
 DR WPI: 90-224494/29.
 PT New non-toxic, immunogenic mutants of pneumolysin - useful in
 PT protective vaccines against Streptococcus pneumoniae, and DNA
 PT sequences encoding them.
 PS Disclosures: P: English.
 CC Vaccines are non-toxic and antigenic to wild type pneumolysin,
 CC making them useful in vaccination against pneumonia and associated
 CC meningitis, bacteraemia etc.
 SQ Sequence 471 AA:

Query Match 94.5%; Score 3056; DB 1; Length 471;
 Best Local Similarity 94.5%; Pred. No. 3,57e-257;

Matches 445; Conservative 6; Mismatches 20; Indels 0; Gaps 0;

Db 1 mankavndfilamnydkkllthggesienrfikegnqlpefviwtekrkststotsdi 60
 |||||
 Qy 1 MANKAVNDFILAMNYDKKLLTHGGESEIENRFKEGNQLPEFVWTEKRKRSLSSTNSDI 60
 Db 61 svratdstrlypgalllvdeetlennptllavdrpmtystldlpglaasdsflgvedpsn 120
 |||||
 Qy 61 XVRATDSRTPGALLVDEETLKENNPILLAVDRAPMTYSTLDLPGLASDSFLGVEDPSN 120
 Db 121 ssrvsgavndllakwbdyqgvnvparymtektahsmeglkvfkgsdfektnsldidf 180
 |||||
 Qy 121 SSRVSGAVNDLLAKWBDYGOVNNVPARXQYKTRAHSMEOULKVKGSDFEKTSGLDIDF 180
 Db 181 nswhsgekgqivnfgkqlyyrvsvdavnpgdfqdvteedlkqrgiseprrlyisv 240
 |||||
 Qy 181 NSVHSEKQIOIVNKKOITVSDAVKKNPGDFQDVTVEEDLKORGISAEPLVYISXV 240
 Db 241 aygrgyvllkettssdeveaafaelikgvkvaoptekqjldntevkavllggdpsga 300
 |||||

Qy 241 AYGRGYVLLKETTSSXSEVEAFAELIKGVKVAOPTEKQJLDNTEVKAVLLGGDPSSGA 300
 Db 301 rvtgkvdmvedlqgesrftadhpglpisytstfirdnvaatfgnstdyvetkayrn 360
 |||||
 Qy 301 RVTGKVDMEVDLQGESRFTADHPGLPISTYTSFLRDNVATFQNSTDYETVKAYRN 360
 Db 361 gdlldhsagayvaayitfdelsfnhgkveitpkafdrngqdltahtstipkgnvrn 420
 |||||
 Qy 361 GDLLDHSGAYVAAYITWNLSDYHOGKEVITPKAMDNGDULTAHTSTIPKGNVRN 420
 Db 421 lsvkirectglawemwrtvyecktdplvrkrtstsgwtltpyvedkveend 471
 |||||
 Qy 421 LSVKIRECTGLAWEMWRTYETKTDPLVRKRRTISMGTTLPYQVEDKVEEND 471

RESULT 3

ID R33842 standard; protein: 480 AA.
 AC R33842:
 DE 15-JUL-1993 (first entry)
 DE Streptolysin O variant msLO.3/6.
 KW SLO; soluble; haemolytic activity; wild type; anti-SLO antibodies;
 OS Streptococcus pyogenes; ASO.
 PN WO9305155-A.
 PD 18-MAR-1993.
 PE 03-AUG-1992; U06380.
 PR 30-AUG-1991; US-752428.
 PA (BECK) BECKMAN INSTR INC.
 PI Adams CW;
 PI WPI: 93-100979/12.
 DR N-PSDB: Q38287.
 PT Streptolysin O variants produced by recombinant DNA technology -
 PT having no haemolytic activity and recognised by wild-type
 PT anti-streptolysin O antibodies, useful in diagnosis of
 PT Streptococcus pyogenes infection
 PS Claim 17: Fig 2: 80pp: English.
 CC The sequence represents a soluble variant of Streptolysin O (SLO)
 CC having no haemolytic activity. The variant SLO can specifically
 CC recognise wild-type anti-SLO antibodies (ASO) but has no haemolytic
 CC activity. It can be used for the determination of previous and/or
 CC current infection by Streptococcus pyogenes. It enables ASO assays
 CC to become commercially viable and avoids the risk of handling haemo-
 CC lytically active S. pyogenes wild-type SLO.
 SQ Sequence 480 AA:

Query Match 43.8%; Score 1416; DB 6; Length 480;
 Best Local Similarity 42.0%; Pred. No. 5,45e-111;

Matches 192; Conservative 104; Mismatches 160; Indels 1; Gaps 1;

Db 18 indkyslnpalevlakngetenfvpkegyvkkadkfivierkkknlttprvdisids 77
 |||||
 Qy 6 VNDFILAMNYDKKLLTHGGESEIENRFKEGNQLPEFVWTEKRKRSLSSTNSDI XVTAT 65
 Db 78 vtdtlypaalqankgtentkpkdavnrtknpqkibldpgmddkt-vevndptyanst 136
 |||||
 Qy 66 NDSRTPGALLVDEETLKENNPILLAVDRAPMTYSTLDLPGLASDSFLGVEDPSNSVVG 125
 Db 137 aichlvqwhnysggnltlpartqytksmysksglaealnvnskllgltlqdfksisk 196
 |||||
 Qy 126 AVNDILAKMHODYGOVNNVPARXQYKTRAHSMEOULKVKGSDFEKTSGLDIDFNSVHS 185
 Db 197 gekwmaiaaykqifvtsaanlpmpnadvfksvtfkelqtkygsneaprlfysnvaqrt 256
 |||||
 Qy 186 GEKOIOIVNKKOITVSDAVKKNPGDFQDVTVEEDLKORGISAEPLVYISXVAYGQ 245
 Db 257 vfvkieteskendeaafaelikgvkvaoptekqjldntevkavllggdpsga 316
 |||||
 Qy 246 VYKLETTSSXSEVEAFAELIKGVKVAOPTEKQJLDNTEVKAVLLGGDPSSGARVVG 305
 Db 317 dfdvlnrvkdnatfsrknpaipysvtfknkklagvnnarteyetsteytsklnl 376
 |||||
 Qy 306 KYDWMVEDLQGESRFTADHPGLPISTYTSFLRDNVATFQNSTDYETVKAYRN DDL 365

Db 377 shgagayvaqyelwdeinyddkqkveitkrrwdnmysktsfvtvlpjgansnrima 436
 QY 366 DHSGAYVAQYIITWNLSTYDHQGEKVLTPKAWDRNGODLTAHFTTSLPLKGNVNLISVKI 425
 Db 437 rectglawewrkylderdvklskelnvngstlsp 473
 QY 426 RECTGLAWEMWRKYERKIDPLVRRKRITISWGTLLP 462

RESULT 4
 ID R33841 standard; Protein; 480 AA.
 AC R33841;
 DT 08-JUL-1993 (first entry)
 DE Soluble streptolysin O variant.
 KW Recombinant; haemolytic activity; immunodiagnostic activity; ss.
 OS Streptococcus pyogenes.
 PN W09J05156-A.
 PD 18-MAR-1993.
 PF 31-JUL-1992; 006398.
 PR 30-AUG-1991; US-752429.
 PA (BECI) BECKMAN INSTR INC.
 PI Adams CW, Wang EY;
 DR WPI: 93-100980/12.
 DR N-PSDB; 038286.
 PT Deriv. of streptolysin O with haemolytic activity - used in
 PT immuno-diagnostic assays which rely on haemolytic activity of
 PS wild-type Streptolysin O
 PS Claim 7; Fig 2; 60pp; English.
 CC The sequence is that of a soluble deriv. of Streptolysin O
 CC with haemolytic properties which was obtd. from Streptococcus pyogenes
 CC genomic library clone rSO.3. It is useful in immunodiagnostic assays
 CC which rely upon, e.g. the haemolytic properties of wild type
 CC Streptolysin O. Recombinant streptolysin O is obtd. more cheaply
 CC than purified streptolysin O obtd. from Streptococcus pyogenes.
 SQ Sequence 480 AA;

Query Match 43.28; Score 1398; DB 6; Length 480;
 Best Local Similarity 41.6%; Pred. No. 2.13e-109;

Matches 190; Conservative 104; Mismatches 162; Indels 1; Gaps 1;

Db 18 indkylslmynelevlakngetienvfpekgvkkadkfivierkkknitpvdslsids 77
 QY 6 VNDFTLAMYDKKKLLTHOGESIENRFXKGNQLPDEFVYXERKKSLSINTSDIXVTAT 65
 Db 78 vdrtypaalqiankgfienkpdavvtrknpqknhidipmgdkat-vevndptyanvt 136
 QY 66 NDSRLYPGALLVYDETLKXNNPTLLAVDRAPMTYSIDLPGLASSDSFLOVEDPSNSVVG 125
 Db 137 aidnlvngndhnysgnltipartqytsesmyksqieaalnvnskildgtlgidfksisk 196
 QY 126 AVNDLLAKWHQDYGQVNNVPAKQYERKTAHSMEOIKVFCGSDFEKTSIDIDFNSVHS 185
 Db 197 gekkvmiaaykqgifytvsanlpnpadvfdkvtfkelgrkyvneapplfvsnyagrt 256
 QY 186 GEKOIOIVNKKQIYTVSDAVANKPGDVFQDTYVEDLKQKGISMERLVYISXVAHQ 245
 Db 257 vfvkletsksndveaafsaalkgtdvktngkysdileensftavvlggdaaehkvtk 316
 QY 246 VYKLETTSSXSEVAEALIKGVKAPQTEWKOILDNTEVKAVALIGDPPSSGARVVG 305
 Db 317 dfdvtrnykdnatfsrknpayisvtsvflknkknlaavmrteyetsgsknl 376
 QY 306 KVDWVEDLIOGSRFTADHPGLPISTTSFLRDNVVATFQNSTDYVEKTVAYRNNGDLL 365
 Db 377 shgagayvaqyelwdeinyddkqkveitkrrwdnmysktsfvtvlpjgansnrima 436
 QY 366 DHSGAYVAQYIITWNLSTYDHQGEKVLTPKAWDRNGODLTAHFTTSLPLKGNVNLISVKI 425
 Db 437 rectglawewrkylderdvklskelnvngstlsp 473
 QY 426 RECTGLAWEMWRKYERKIDPLVRRKRITISWGTLLP 462

RESULT 5
 ID R10376 standard; Protein; 571 AA.
 AC R10376;
 DT 05-APR-1991 (first entry)
 DE Streptolysin O derivative.
 KW SLO.
 OS Streptococcus pyogenes.
 PN GB2233977-A.
 PD 23-JAN-1991.
 PF 04-JAN-1989; 000107.
 PR 04-JAN-1989; GB-000107.
 PA (KEHO) KEHOE M.
 PI (PINK) PINKNEY M.
 PI Kehoe M, Pinkney M;
 DR WPI: 91-024598/04.
 DR N-PSDB; 010320.
 PT Deriv. of thiol-activated protein streptolysin O - contg. no
 PT cysteine amino acid but retaining cytolitic activity, used for
 PT detecting antibodies in samples
 PS Disclosure; Fig 2; 9pp; English.
 CC The SLO derivative contains no cysteine residues, with cytolitic
 CC activity giving resistance to inactivation by oxidation or
 CC thiol groups. Abs raised to the SLO may be used in the decton and
 CC diagnosis of Streptococcus pyogenes infection.
 SQ Sequence 571 AA;

Query Match 43.0%; Score 1391; DB 2; Length 571;

Best Local Similarity 41.6%; Pred. No. 8.88e-109;

Matches 190; Conservative 104; Mismatches 162; Indels 1; Gaps 1;

Db 109 indkylslmynelevlakngetienvfpekgvkkadkfivierkkknitpvdslsids 168
 QY 6 VNDFTLAMYDKKKLLTHOGESIENRFXKGNQLPDEFVYXERKKSLSINTSDIXVTAT 65
 Db 169 vdrtypaalqiankgfienkpdavvtrknpqknhidipmgdkat-vevndptyanvt 227
 QY 66 NDSRLYPGALLVYDETLKXNNPTLLAVDRAPMTYSIDLPGLASSDSFLOVEDPSNSVVG 125
 Db 228 aidnlvngndhnysgnltipartqytsesmyksqieaalnvnskildgtlgidfksisk 287
 QY 126 AVNDLLAKWHQDYGQVNNVPAKQYERKTAHSMEOIKVFCGSDFEKTSIDIDFNSVHS 185
 Db 288 gekkvmiaaykqgifytvsanlpnpadvfdkvtfkelgrkyvneapplfvsnyagrt 347
 QY 186 GEKOIOIVNKKQIYTVSDAVANKPGDVFQDTYVEDLKQKGISMERLVYISXVAHQ 245
 Db 348 vfvkletsksndveaafsaalkgtdvktngkysdileensftavvlggdaaehkvtk 407
 QY 246 VYKLETTSSXSEVAEALIKGVKAPQTEWKOILDNTEVKAVALIGDPPSSGARVVG 305
 Db 408 dfdvtrnykdnatfsrknpayisvtsvflknkknlaavmrteyetsgsknl 467
 QY 306 KVDWVEDLIOGSRFTADHPGLPISTTSFLRDNVVATFQNSTDYVEKTVAYRNNGDLL 365
 Db 468 shgagayvaqyelwdeinyddkqkveitkrrwdnmysktsfvtvlpjgansnrima 527
 QY 366 DHSGAYVAQYIITWNLSTYDHQGEKVLTPKAWDRNGODLTAHFTTSLPLKGNVNLISVKI 425
 Db 528 rectglawewrkylderdvklskelnvngstlsp 564
 QY 426 RECTGLAWEMWRKYERKIDPLVRRKRITISWGTLLP 462

RESULT 6
 ID R06000 standard; Protein; 371 AA.
 AC R06000;
 DT 22-NOV-1990 (first entry)
 DE Bacteriophage lambda PL promoter - streptolysin O (SLO) fusion
 DE protein of plasmid pmk306.
 KW Streptolysin O; SLO; protease; bacteriophage lambda; ds.
 OS Streptococcus pyogenes.
 FT Key Location/Qualifiers
 FT cleavage_site 33..34

PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berglindh OT, Smith D, Mellgaard BL.
 DR WPI: 97-052306/05.
 DR N-PSDB: T68006.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PS Claim 56; Page 1166-1167; 1481pp; English.
 CC The present sequence is a Helicobacter pylori transporter protein.
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 CC and the predicted coding regions defined by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SQ Sequence 667 AA;

Query Match 3.1%; Score 101; DB 22; Length 667;
 Best Local Similarity 26.9%; Pred. No. 3,16e+01;
 Matches 14; Conservative 15; Mismatches 22; Indels 1; Gaps 1;

DB 533 lkgldfvqvgfkgvgrantstcda-agiqlmetyelkxqnewkeklxykevr 583
 Oy 237 ISXVAYGKQVYLKLETTTSXSEVEAFALIKGVAVAPQTEMKQILDWTEVK 288

RESULT 15
 ID R41781 standard; Protein; 458 AA.
 AC R41781;
 DE 28-MAR-1994 (first entry)
 KW Rabbit HBI protein which binds hsp90.
 KW steroid hormone receptor; immunosuppressor; FK506; binding protein;
 KW FKBP; heat shock protein; hsp90; chaperone protein;
 OS rotamase activity; ss.
 KM Oryctolagus cuniculus.
 FH Key
 FT Location/Qualifiers
 FT domain
 FT 32..138
 FT /label= HBI-I
 FT /note= "immunophilin domain; has 49 per cent
 FT identity with FKBP without introducing
 FT insertions or deletions into the sequence"
 FT 139..148
 FT /label= hinge
 FT 149..253
 FT /label= HBI-II
 FT /note= "immunophilin domain; has 28 per cent
 FT identity with FKBP and comprises 4 zones of
 FT insertions or deletions"
 FT 254..267
 FT /label= hinge
 FT 268..372
 FT /label= HBI-III
 FT /note= "globular domain"
 FT 373..458
 FT /label= C-terminal
 FT /note= "CAM"
 FT 33..39
 FT /label= A1
 FT /note= "antiparallel beta-pleated sheet"
 FT 53..61
 FT /label= A4
 FT /note= "antiparallel beta-pleated sheet"
 FT 66..68
 FT /label= A5
 FT /note= "antiparallel beta-pleated sheet"

FT region
 FT /label= beta
 FT /note= "supplementary beta-pleated sheet"
 FT 88..97
 FT /note= "alpha-helix"
 FT 103..106
 FT /label= A2
 FT /note= "antiparallel beta-pleated sheet"
 FT 128..137
 FT /label= A3
 FT /note= "antiparallel beta-pleated sheet"
 FT 150..156
 FT /label= A1
 FT /note= "antiparallel beta-pleated sheet"
 FT 170..178
 FT /label= A4
 FT /note= "antiparallel beta-pleated sheet"
 FT 182..184
 FT /label= A5
 FT /note= "antiparallel beta-pleated sheet"
 FT 188..192
 FT /label= beta
 FT /note= "supplementary beta-pleated sheet"
 FT 202..211
 FT /note= "alpha-helix"
 FT 217..220
 FT /label= A2
 FT /note= "antiparallel beta-pleated sheet"
 FT 243..252
 FT /label= A3
 FT /note= "antiparallel beta-pleated sheet"
 FT 269..275
 FT /label= A1
 FT /note= "antiparallel beta-pleated sheet"
 FT 289..297
 FT /label= A4
 FT /note= "antiparallel beta-pleated sheet"
 FT 301..303
 FT /label= A5
 FT /note= "antiparallel beta-pleated sheet"
 FT 307..311
 FT /label= beta
 FT /note= "supplementary beta-pleated sheet"
 FT 321..330
 FT /note= "alpha-helix"
 FT 335..338
 FT /label= A2
 FT /note= "antiparallel beta-pleated sheet"
 FT 362..371
 FT /label= A3
 FT /note= "antiparallel beta-pleated sheet"
 FT 399..414
 FT /note= "site of interaction with calmodulin"
 FT 441..458
 FT /note= "used to generate polyclonal antibodies"
 FT region
 FT binding_site
 FT 16-SEP-1993.
 FT 04-MAR-1993; F00219.
 FT (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 FT Baulieu E, Callebaut I, Chambraud B, Ledreau M, Massol N;
 FT WPI: 93-303460/38.
 FT N-PSDB: Q48515.
 FT New protein forming complex with heat shock protein - also
 FT binding immuno-suppressors, etc., and corresp. nucleic acid,
 FT antibodies etc., useful e.g. for detecting tumours, treating
 FT auto-immune disease, etc.
 FT Claim 1, Fig 1; 43pp; French.
 FT The HBI protein is able to complex chaperone protein hsp90, even
 FT when the chaperone protein is part of a hetero-oligomer with other
 FT proteins. Hsp90 can bind to steroid hormone receptors, vitamin D,
 FT and Tyrosine Kinases of viral oncogenes; HBI is thus useful for the

Fri Sep 3 14:48:31 1999

US-09-120-044-3-COPY.rag

Page 8

CC study, prevention or treatment of diseases associated with
CC dysfunction of proteins which form complexes with hsp90, e.g.,
CC rickets, cancer, dioxin poisoning or autoimmune disease. HBI can
CC also be used in immunosuppressor research.
SQ Sequence 458 AA;

CC dysfunction of proteins which form complexes with hsp90, e.g.

CC also be used in immunosuppressor research.

sq Sequence 458 AA;

Query Match	3.08; Score 96; DB 8; Length 458;
-------------	-----------------------------------

Query Match	3.0%;	Score 96;	DB 8;	Length 458;
Best Local Similarity	22.2%;	Pred. No. 6.61e+01;		
Matches	20;	Conservative	49;	Indels 2;
		Mismatches		Gaps 2;

Db 221 lkpsyaftgnagkekfqipppaelkyevhlksfekakesw-emsseeklegsaivkergtv 279

176 LDIDENSVHSEKQIQVNKKQYYTVSVDAVKNPGDVQDTVIEDLKQGGISAEPLV 235

Db 280 yfkeqky-kgalqykvivswleyessfss 308
 | . | . | . | . | . | . | . | .
 | . | . | . | . | . | . | . | .

QY 236 YISXVAYGRÖVYLKLETTSSXSXEVEAEFA 265

```
Search completed: Fri Sep 3 13:13:36 1999
Job time : 110 secs.
```

```
Search completed: Fri Sep 3 13:13:36 1999
Job time : 110 secs.
```

Job time : 110 secs.

Db 121 SSVGAANDLAKMHQDYGVYNNVPAKMOYEKITAHSMEQLKVKFGSDPEKIGNSLIDIF 180
121 SSVGAANDLAKMHQDYGVYNNVPAKMOYEKITAHSMEQLKVKFGSDPEKIGNSLIDIF 180
Qy 181 NSVHSGEKQIOIVFKQIYITVSDAVKPNQDVFQDITYVEDLKQKISAEPLVYISXV 240
181 NSVHSGEKQIOIVFKQIYITVSDAVKPNQDVFQDITYVEDLKQKISAEPLVYISXV 240
Qy 241 AYGVQVYKLETTESKDEVEAFALIKGVKVAPOTEMKOILNTEKAVIILGDPSSGA 300
241 AYGVQVYKLETTESKDEVEAFALIKGVKVAPOTEMKOILNTEKAVIILGDPSSGA 300
Qy 301 RVATGVKDMVEDLQEGSRFTADHPGLPISTYTSFLRDNVVAFFQNSTDYETKVTAYRN 360
301 RVATGVKDMVEDLQEGSRFTADHPGLPISTYTSFLRDNVVAFFQNSTDYETKVTAYRN 360
Db 361 GDLLDHSGAVVAQYIYITWELSYDHQKEVLTFRKAMDNGODLTAFTTISIPKGNVRN 420
361 GDLLDHSGAVVAQYIYITWELSYDHQKEVLTFRKAMDNGODLTAFTTISIPKGNVRN 420
Qy 421 LSVKIRECTGLAMEMWRTVYEKTDLPVRRKRTISMGTTIYPQVEDKVEN 471
421 LSVKIRECTGLAMEMWRTVYEKTDLPVRRKRTISMGTTIYPQVEDKVEN 471
Qy 471 LSVKIRECTGLAMEMWRTVYEKTDLPVRRKRTISMGTTIYPQVEDKVEN 471
471 LSVKIRECTGLAMEMWRTVYEKTDLPVRRKRTISMGTTIYPQVEDKVEN 471
RESULT 2 B43577 #type complete
ENTRY perfringolysin O precursor - Clostridium perfringens
TITLE #formal_name Clostridium perfringens
ORGANISM 03-Feb-1993 #sequence_revision 03-Mar-1993 #text_change
DATE 13-Sep-1998
ACCESSIONS B43577; A34951; A60922
REFERENCE A43577
#authors Shimizu, T.; Okabe, A.; Minami, J.; Hayashi, H.
#journal Infect. Immun. (1991) 59:137-142
#title An upstream regulatory sequence stimulates expression of the
perfringolysin O gene of Clostridium perfringens.
#cross-references MUID:91099951
#accession B43577
#molecule_type DNA
#residues 1-500 #label SHI
#cross-references GB:M81080; NID:9144885; PID:9144886
#note translation of the nucleotide sequence is not complete
REFERENCE A34951
#authors Tweten, R.K.
#journal Infect. Immun. (1988) 56:3235-3240
#title Nucleotide sequence of the gene for perfringolysin O
(theta-toxin) from Clostridium perfringens: significant
homology with the genes for streptolysin O and pneumolysin.
#cross-references MUID:89032623
#accession A34951
#molecule_type DNA
#residues 1-125, 'EA', 129-500 #label TWE
#cross-references GB:M36704; NID:9144883; PID:9144884
REFERENCE A60922
#authors Tweten, R.K.
#journal Infect. Immun. (1988) 56:3228-3234
#title Cloning and expression in Escherichia coli of the
perfringolysin O (theta-toxin) gene from Clostridium
perfringens and characterization of the gene product.
#cross-references MUID:89032622
#accession A60922
#molecule_type protein
#residues 29-45 #label TW2
#experimental_source ATCC 13124
GENETICS pfo; pfoA
#gene
FEATURE 1-28 #domain signal sequence #status predicted #label SIGV
29-500 #product perfringolysin O #status experimental #label
MAR
SUMMARY #length 500 #molecular-weight 55799 #checksum 4821

Query Match 49.5%; Score 1600; DB 2; Length 500;
Best Local Similarity 46.5%; Pred. No. 2,15e-270;
Matches 214; Conservative 105; Mismatches 140; Indels 1; Gaps 1;
Db 35 NOSIDSGISSLSTYNNRNEVLAANGKISFVPEKSKKGNKFNIVYERKRRLTSPVDISI 94
35 NOSIDSGISSLSTYNNRNEVLAANGKISFVPEKSKKGNKFNIVYERKRRLTSPVDISI 94
Qy 3 NKAIVNDFLIANNKKKLLTHQGESIENRFKEGNOQJDEFEVXERRKRLSTNTSDIXV 62
3 NKAIVNDFLIANNKKKLLTHQGESIENRFKEGNOQJDEFEVXERRKRLSTNTSDIXV 62
Db 95 IDSVNDRTPYPCALODADKAFENRPTILMYKRKPININIDPLGKJGNS-1KYDDPTYGK 153
95 IDSVNDRTPYPCALODADKAFENRPTILMYKRKPININIDPLGKJGNS-1KYDDPTYGK 153
Qy 63 TATNDSRLYPCALLVDETLKENNPITLADRAMYTSIDLPGLAASDPSLOYDEPENSS 122
63 TATNDSRLYPCALLVDETLKENNPITLADRAMYTSIDLPGLAASDPSLOYDEPENSS 122
Db 123 VSGAIDELVSKWNEKYSSTHTLPARTQYSESMSYKSKQISSALNNAKVLNLSIGVDENA 213
123 VSGAIDELVSKWNEKYSSTHTLPARTQYSESMSYKSKQISSALNNAKVLNLSIGVDENA 213
Qy 123 VSGAANDLAKMHQDYGVYNNVPAKMOYEKITAHSMEQLKVKFGSDPEKIGNSLIDIFNS 182
123 VSGAANDLAKMHQDYGVYNNVPAKMOYEKITAHSMEQLKVKFGSDPEKIGNSLIDIFNS 182
Db 214 VANNKKVMILAYKQIEFTVSADLPKPNSDLPDSDSVFNOLKQKGVNENPPLMVSVAV 273
214 VANNKKVMILAYKQIEFTVSADLPKPNSDLPDSDSVFNOLKQKGVNENPPLMVSVAV 273
Qy 183 VHSGEKQIOIVNKKQIYITVSDAVKPNQDVFQDITYVEDLKQKISAEPLVYISXV 242
183 VHSGEKQIOIVNKKQIYITVSDAVKPNQDVFQDITYVEDLKQKISAEPLVYISXV 242
Db 274 GRTIVYKLETTSSKDVQAAFKALIKNDIKNSQOYKDIYENSFTAVVLGDAQEHNKV 333
274 GRTIVYKLETTSSKDVQAAFKALIKNDIKNSQOYKDIYENSFTAVVLGDAQEHNKV 333
Qy 243 GRQVYKLETTSSKDEVEAFALIKGVKVAPOTEMKOILNTEKAVIILGDPSSGARV 302
243 GRQVYKLETTSSKDEVEAFALIKGVKVAPOTEMKOILNTEKAVIILGDPSSGARV 302
Db 334 YTKDFEIRKVIKDNATESTKPNAPVISTYSVFLKDNVAAVHNKTDYIETTSTESKKG 393
334 YTKDFEIRKVIKDNATESTKPNAPVISTYSVFLKDNVAAVHNKTDYIETTSTESKKG 393
Qy 303 YVGVKDMVEDLQEGSRFTADHPGLPISTYTSFLRDNVVAFFQNSTDYETKVTAYRN 362
303 YVGVKDMVEDLQEGSRFTADHPGLPISTYTSFLRDNVVAFFQNSTDYETKVTAYRN 362
Db 394 INLDHSGAVVAQYIYITWELSYDHQKEVLTFRKAMDNGODLTAFTTISIPKGNVRN 453
394 INLDHSGAVVAQYIYITWELSYDHQKEVLTFRKAMDNGODLTAFTTISIPKGNVRN 453
Qy 363 LLLDHSGAVVAQYIYITWELSYDHQKEVLTFRKAMDNGODLTAFTTISIPKGNVRN 422
363 LLLDHSGAVVAQYIYITWELSYDHQKEVLTFRKAMDNGODLTAFTTISIPKGNVRN 422
Db 454 IKARECTGLAMEMWRTVYEKTDLPVRRKRTISMGTTIYP 493
454 IKARECTGLAMEMWRTVYEKTDLPVRRKRTISMGTTIYP 493
Qy 423 VKIRECTGLAMEMWRTVYEKTDLPVRRKRTISMGTTIYP 462
423 VKIRECTGLAMEMWRTVYEKTDLPVRRKRTISMGTTIYP 462
RESULT 3 A43505 #type complete
ENTRY listeriolysin O precursor - Listeria monocytogenes
TITLE #formal_name Listeria monocytogenes
ORGANISM 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change
DATE 17-Mar-1999
ACCESSIONS A43505; S05306; A47606; S12400; A61079
REFERENCE A43505
#authors Mengaud, J.; Vicente, M.F.; Chenevert, J.; Pereira, J.M.;
Geoffroy, C.; Gicquel-Sanzey, B.; Baquero, F.; Perez-Diaz,
J.C.; Cossart, P.
#journal Infect. Immun. (1988) 56:766-772
#title Expression in Escherichia coli and sequence analysis of the
listeriolysin O determinant of Listeria monocytogenes.
#cross-references MUID:88153053
#accession A43505
#status preliminary
#molecule_type DNA
#residues 1-529 #label MEN
#cross-references GB:M24199; NID:9149652; PID:9149653
#note this sequence is derived from a strongly hemolytic
strain, serotype 1/2c
REFERENCE S05306
#authors Domann, E.; Chakraborty, T.
#journal Nucleic Acids Res. (1989) 17:6406
#title Nucleotide sequence of the listeriolysin gene from a Listeria
monocytogenes serotype 1/2a strain.
#cross-references MUID:89366684
#accession S05306
#molecule_type DNA
#residues 1-529 #label DOM
#cross-references EMBL:X15127; NID:944106; PID:944107
#experimental_source strain Egd
#note this sequence is derived from a weakly hemolytic strain,
serotype 1/2a
REFERENCE A47606

[illegible]

ORGANISM	#formal_name	Listeria monocytogenes
DATE	12067)	
DATE	variety	strain 12067
DATE	22-Nov-1993	#sequence_revision 10-Nov-1995
DATE	13-Sep-1998	#text_change
ACCESSIONS	S24231	
REFERENCE	S24230	
#authors	Rasmussen, O.F.; Beck, T.; Olsen, J.E.; Dons, L.; Rossen, L.	
#journal	Infect. Immun. (1991) 59:3945-3951	
#title	Listeria monocytogenes isolates can be classified into two major types according to the sequence of the listeriolysin gene.	
#cross-references	EMBL:292040062	
#accession	S24231	
#status	nucleic acid sequence not shown; translation not shown	
##molecule_type	DNA	
##residues	1-529	##label RAS
##cross-references	EMBL:X60035; NID:g44110; PID:g44112	
##experimental_source	strain 12067, serotype 4b	
##note	the nucleotide sequence was submitted to the EMBL data library, June 1991	
GENETICS		
##gene	lisa	
CLASSIFICATION	##superfamily dipeptide transport protein	
KEYWORDS	virulence factor	
FEATURE		
1-25	##domain signal sequence	#status predicted
26-529	##product listeriolysin	#status predicted
SUMMARY	##length 529	##molecular_weight 58687
Query Match	47.5%	Score 1536; DB 2; Length 529;
Best Local Similarity	42.7%;	Pred. No. 4,066-258;
Matches	199; Conservative 117;	Mismatches 149; Indels 1; Gaps 1
Db	61 IDKTYOGLDYNNNNVLYHGDAVTWVPRKQYKQGENEYVEKKKKKSIQNNADIQVYNA	120
Qy	6 VADFLAANNYKKKLLTHQGESIERFRKBEQNLDPDEVVAXERKKRSLSTSDIXYAT	65
Db	121 ISSLTYPGALVKANSELVENOPDVLVPRKDSLTLSIDLPGMTNODNKIVYNAKTSVYNN	180
Qy	66 NDSRLYPGALLVDETLKXENNPITLLAVDRAPMTYSIDLPLGLASSDSFLOVEDPSSVVG	125
Db	181 AVNTLVERNEKYAQAIPYNSAKIDYDDEMAVSESQILAKCTAFKAVNSLNVEGALS	240
Qy	126 AVNDLAKHODYGVN-NVPAKQYKERTASHMQLKVKVSGSEFKTSGSLDIDENSVAH	184
Db	241 EKGMOEVSFQIYYNNVNPETPRSPFFGKAVKQDOLAGVAENRPAVYISVAVGR	300
Qy	185 SEKQIQIYNKQIYYTYSVDVAKKPGVFDQTYVEDLKQGISAEKRPVYISVAVGR	244
Db	301 QVYLKLTINSHSTYKKAADFVAASGVSGDVELTINIKSSFFRAVYIGSGAKDEVOIID	360
Qy	245 QVYLKLETTSSKXEVEAFALIKGVKAPQEMKQIIDNEVKAIVLIGDPPSGARVYT	304
Db	361 GNLGDLRLIKGAFNRETPEVPPIATYTNPLKQDELAVIKNSGYITTSKATYDGIN	420
Qy	305 GYVDWVEDIIOGGSFRTADHPCLPSYTTSTFLRDVAVTAFONSDYVETKYTAARNGDL	364
Db	421 IDHSGGVQAFINSDMEIYDEGNEIYOHKWSSENNKSKLAHFSSLYLPGMARINWY	480
Qy	365 LDHSGAYVAQYITTNELSDYHOGKEVLLPRKMRDNGDRLAHFTTSLPLKGNVRLSVK	424
Db	481 AECTGLAEMWRVYIDRNLPLVKNRNIISWGLTYPKYSNSVDN	526
Qy	425 IRECTGLAEMWRVYERKTDLPVYKRTISWGTLLYPOVEDKVEN	470
RESULT	5	
ENTRY	S22341	#type complete
TITLE	livanolysin precursor - Listeria ivanovii	
ORGANISM	07-Apr-1994	#formal_name Listeria ivanovii
DATE	21-Aug-1998	#sequence_revision 07-Apr-1994
		#text_change

ACCESIONS	S22341; S36683
REFERENCE	S22340
#authors	Haas, A.; Dumbeky, M.; Kreft, J.
#journal	Biochim. Biophys. Acta (1992) 1130:81-84
#title	Listeriolysin genes: complete sequence of <i>ilo</i> from <i>Listeria</i> Ivanovii and of <i>iso</i> from <i>Listeria seeligeri</i> .
#cross-references	MUID:92182018
#accession	S22341
##molecule_type	DNA
##residues	1-528 ##label HAS
##cross-references	EMBL:X60461
##note	the authors translated the codon ACA for residue 331 as Val
REFERENCE	S36683
#authors	Kreft, J.
#submission	submitted to the EMBL Data Library, July 1991
#accession	S36683
##molecule_type	DNA
##residues	1-319, 'T', 321-528 ##label KRE
##cross-references	EMBL:X60461
GENETICS	
#gene	<i>ilo</i>
CLASSIFICATION	#superfamily dipeptide transport protein
FEATURE	
1-24	
25-528	#domain signal sequence #status predicted #label SIG\
SUMMARY	#product <i>ilv</i> #status predicted #label MAT
Query Match	47.2%; Score 1526; DB 2; Length 528;
Best Local Similarity	42.9%; Pred. No. 3.35e-256;
Matches	200; Conservative 118; Mismatches 147; Indels 1; Gaps 1;
Db	60 IDQYIOGLDYDKRNILVYDGEAVNVNPKAKYKGGNOIVYERKKKSINONNADIOVINS 119
QY	6 VNDFTLANVYDKKLLTHQGESIENRF*KEGNQLPDERVYERKKRSLSTNSDIXATAT 65
Db	120 LASLTYPALYKANSELVENOPDYLPYKRDSYTLSTIDPGAVNHDNEIVQNAATKSNIND 179
QY	66 NDSRLYPALLLVDETLKENNPFTLLADRAMPTYSIDLPGCLASSDSFLQVEDPENSVRG 125
Db	180 GVNTLVYDMMNNKYSEEEYNISAKIDYDQEMAYSSQIYAKGAFAKAVNNSLAWFEAIS 239
QY	126 AVNDLLAMWHDQYQVNVNPARQYKEITIHSHQQLKAFGSGFEETGSLSDIDENSVH 184
Db	240 EGYQVEEYINFKQIYIVVNVNEPTSPSRFEKSVTKENLQALGVNAENPPAYISSVAYGR 299
QY	185 SGEKQIOIVNNKQIYIVSVDAVNVPGVDFQDTVEDLKQRGISARPLVYISSVAYGR 244
Db	300 DIFVKLSTSSSTRYKKAFAADAFKSGKSGTLENTIIQNASFRAYVYGSADVEIID 359
QY	245 QYVLTLETTSSKSEVEAAFEALINGVAYAPOTEPKQILINDTEKAAVILGDPSSGARVYT 304
Db	360 GDSLKLDIILKOGANFEDKKNPGVPIAATYTNFLKQNOIAVYKNNSEYIETTSKAYSDGKIN 419
QY	305 GKVMVVEDLLIOEGSRFTADHBPGLISYTTSPFLRDNVAVATFQNSDYETKVTAYVRGDDL 364
Db	420 LDHSGAVYARVNVWDESYDANENVEYEHKKWSBENKDKLAHTTSTIYIPGAKRNINIH 479
QY	365 LDHSGAVYAYQIYITWNEISIDHQSKEVLTPAPAMRNQODLLAHTTSTIPLKGVNRLSYK 424
Db	480 AKECTGLAMEMWRTVVDNRNLPLVKNRNVCITWGTTLTPAYSDVEVDN 525
QY	425 IRECTGLAMEMWRTVYEKTDLPVLRKRTISITWGTTLTPQVEDKAYEN 470
RESULT	6
ENTRY	S22340
TITLE	seeligeriolysin - <i>Listeria seeligeri</i>
ORGANISM	#formal_name <i>Listeria seeligeri</i>
DATE	22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 21-Aug-1998
ACCESIONS	S22340
REFERENCE	S22340

Query Match	Best Local Similarity	Matches	Score	DB 2:	Length	530:
188	45.7%	42.58	1477	DB 2:	Length	530:
187	42.58	42.58	1477	DB 2:	Length	530:
186	42.58	42.58	1477	DB 2:	Length	530:
185	42.58	42.58	1477	DB 2:	Length	530:
184	42.58	42.58	1477	DB 2:	Length	530:
183	42.58	42.58	1477	DB 2:	Length	530:
182	42.58	42.58	1477	DB 2:	Length	530:
181	42.58	42.58	1477	DB 2:	Length	530:
180	42.58	42.58	1477	DB 2:	Length	530:
179	42.58	42.58	1477	DB 2:	Length	530:
178	42.58	42.58	1477	DB 2:	Length	530:
177	42.58	42.58	1477	DB 2:	Length	530:
176	42.58	42.58	1477	DB 2:	Length	530:
175	42.58	42.58	1477	DB 2:	Length	530:
174	42.58	42.58	1477	DB 2:	Length	530:
173	42.58	42.58	1477	DB 2:	Length	530:
172	42.58	42.58	1477	DB 2:	Length	530:
171	42.58	42.58	1477	DB 2:	Length	530:
170	42.58	42.58	1477	DB 2:	Length	530:
169	42.58	42.58	1477	DB 2:	Length	530:
168	42.58	42.58	1477	DB 2:	Length	530:
167	42.58	42.58	1477	DB 2:	Length	530:
166	42.58	42.58	1477	DB 2:	Length	530:
165	42.58	42.58	1477	DB 2:	Length	530:
164	42.58	42.58	1477	DB 2:	Length	530:
163	42.58	42.58	1477	DB 2:	Length	530:
162	42.58	42.58	1477	DB 2:	Length	530:
161	42.58	42.58	1477	DB 2:	Length	530:
160	42.58	42.58	1477	DB 2:	Length	530:
159	42.58	42.58	1477	DB 2:	Length	530:
158	42.58	42.58	1477	DB 2:	Length	530:
157	42.58	42.58	1477	DB 2:	Length	530:
156	42.58	42.58	1477	DB 2:	Length	530:
155	42.58	42.58	1477	DB 2:	Length	530:
154	42.58	42.58	1477	DB 2:	Length	530:
153	42.58	42.58	1477	DB 2:	Length	530:
152	42.58	42.58	1477	DB 2:	Length	530:
151	42.58	42.58	1477	DB 2:	Length	530:
150	42.58	42.58	1477	DB 2:	Length	530:
149	42.58	42.58	1477	DB 2:	Length	530:
148	42.58	42.58	1477	DB 2:	Length	530:
147	42.58	42.58	1477	DB 2:	Length	530:
146	42.58	42.58	1477	DB 2:	Length	530:
145	42.58	42.58	1477	DB 2:	Length	530:
144	42.58	42.58	1477	DB 2:	Length	530:
143	42.58	42.58	1477	DB 2:	Length	530:
142	42.58	42.58	1477	DB 2:	Length	530:
141	42.58	42.58	1477	DB 2:	Length	530:
140	42.58	42.58	1477	DB 2:	Length	530:
139	42.58	42.58	1477	DB 2:	Length	530:
138	42.58	42.58	1477	DB 2:	Length	530:
137	42.58	42.58	1477	DB 2:	Length	530:
136	42.58	42.58	1477	DB 2:	Length	530:
135	42.58	42.58	1477	DB 2:	Length	530:
134	42.58	42.58	1477	DB 2:	Length	530:

Qy	126	AVNDLLAKHNDQYGVGVNNVPARXQYKERTIHAHSNEQJLVKFGSPFEKTSNLDIDENSVHS	185
Db	288	GEKKVIAAAYKQIFTTVSANLPNNPADVEDKSTFEKLOKGYKSNAPLPFSNAYGRT	347
Qy	186	GEKOIIVMKKQIYYTVSDVAKNPGDFODTVVEDLKGKRGISABRPVYISXVAYGRQ	245
Db	348	VFKLETSKSNVDVEAFSAALKGDVTKNGKYSDLIENSFPAAVLGGDAAEHKKVYT	407
Qy	246	VYKLETSKXSEVEAEFALIKGYKAPQTEMKQILDNTEYKAAVLGGDPSSGARVYT	305
Db	408	DFVIRNVIKDNATFSRKNPAPISYTSVFLKNNKLAGVNNRTEYETSTEYTSKINTL	467
Qy	306	KVMDVEDLLQEGSRFLADHPGLPISYTSFLRDVAVTFFQNSIDYETKVTAYRNGDLL	365
Db	468	SHOGAYVAQYEIIMDEINTDDKGVITTKRMDNNWTKSPSYIYPLGANSRNRIRMA	527
Qy	366	DHSGAAVQAQYIYIMWELSYDHGKFEVLTPKAMDNRNGDGLFAHFTTIPLKGVNRLSVYI	425
Db	528	REGCGLAEMWKRVIDERPVTKSRKINNVNISGRSLSP	564
Qy	426	REGCGLAEMWKRIVYEKTDLPVLRKRTISMGITLTP	462
RESULT	9		
ENTRY		I39863	#type fragment
TITLE		hemolysin - Bacillus cereus (fragment)	
ORGANISM		#journal_name Bacillus cereus	
DATE		19-Jul-1996	#sequence_revision 19-Jul-1996
		19-Jul-1996	#text_change
ACCESSIONS		I39863	
REFERENCES		I39863	
authors		Yutsudo, T.; Okumura, K.; Iwasaaki, M.; Hara, A.; Kakitani, S.; Minamide, W.; Igarashi, H.; Hinuma, Y.	
#journal		Infect. Immun. (1994) 62:4000-4004	
#title		The gene encoding a new mitogenic factor in a Streptococcus pyogenes strain is a distributed only in group A streptococci	
#cross-references		UID:94341910	
#accession		I39863	
#status		preliminary; translated from GB/EMBL/DBJ	
#molecule_type		DNA	
#residues		1-485	#label RES
#cross-references		GB:D21270; NID:9418066; PID:9600252	
SUMMARY		#length 485	#checksum 3782
Query Match		41.1%;	Score 1329; DB 2; Length 485;
Best Local Similarity		39.5%;	Pred. No. 1,556-218;
Matches		117;	Mismatches 145; Indels 1; Gaps 1
Db	44	SIDTIGIGNLTNNQOEVLAVANGDKVESYVPKESINSNGKFFVYVYRKHLDORHOSIFRLD	103
Qy	5	AVNDFIILAMYDKKLLTHQGESIENRKEKGNQLPDEFVYXERKKRSLSTNTSDIXAYTA	64
Ddb	104	SVARITPGAVOLANKAPADNOPSLILAKRKPLNLSIDLPKREKT -ITVONPTGYNA	162
Qy	65	TMDSKLPGALVVDDELTKENPFLILANDRAMYTSIDLPELASDSFLQYEDPSSNSVRA	124
Db	163	GAVDLVSTNMEKYSATHTLPARMQYTESMYSKQAQISALUNAVAKYLDNSLINDENAVA	222
Qy	125	GAVNDLLAKHNDQYGVGVNNVPARXQYKERTIHAHSNEQJLVKFGSPFEKTSNLDIDENSVH	184
Db	223	NGEKKVMAAYKQIFTTVSANLPNNPADVEDKSTFEKLOKGYKSNAPLPFSNAYGRT	282
Qy	185	SEKQIOIIVMKKQIYYTVSDVAKNPGDFODTVVEDLKGKRGISABRPVYISXVAYGR	244
Db	283	TYVYKLETSKSNVDVEAFSAALKGDVTKNGKYSDLIENSFPAAVLGGDAAEHKKVYT	342
Qy	245	OYTLKLETSKXSEVEAEFALIKGYKAPQTEMKQILDNTEYKAAVLGGDPSSGARVYT	304
Db	344	KDFENIRNIIIDKNAESFKMAPYISYTSFLKDNATAAVANNNTDIETTTTEYSKAKMT	402
Qy	305	GVMDVEDLLQEGSRFLADHPGLPISYTSFLRDVAVTFFQNSIDYETKVTAYRNGDLL	364

Db	403	LDHGAAYVQAFDVSWMGFEDDQGRKILTHKTHWEGSGKRTAHYVIFPLPNSKNIKTV	462
QY	365	LDHSGAVVAQYITVITWELSTDHQGRKVLTPKAMDNRNGODLTAHFTTSLPKGNVRLTSVK	424
Db	463	ARECTGLAEMWRTI	477
QY	425	IRECTGLAEMWRTV	439
RESULT	10		
ENTRY		s47298	#type complete
TITLE		sulysin - Streptococcus suis	
ORGANISM		#formal_name Streptococcus suis	
DATE		06-Jan-1995	#sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
ACCESSIONS		s47298	
REFERENCE		s47297	
#authors		Segers, R.	
#submission		submitted to the EMBL Data Library, September 1994	
#accession		s47298	
##status		preliminary	
##molecule_type		DNA	
##residues		1-112	#label SEG
##cross-references		EMBL:236908; NID:g535307; PID:g535308	
SUMMARY		#length 112	#molecular_weight 12464 #checksum 9264
Query Match		14.3%; Score 464; DB 2; Length 112;	
Best Local Similarity		50.9%; Pred. No. 2,28e-57;	
Matches		57; Conservative 22; Mismatches 33; Indels 0; Gaps 0	
Db	1	PGVPVSTTFYXXXXPQXLSNSEIETFTSVHSSALTLXSGAIVAKNITGCEVSY	60
QY	325	PGLPSTYITSELRDVAFAFQNSTDYVETKVAIRNGDLLDHSAGVAAQYITWELSLSY	384
Db	61	NEAGEWEKXAMKADKNGVNLISHMSEETIXIPGNAXNLAHYNIQECTGLAEMW	112
QY	385	DHQGEVLPKAMDNRNGODLTAHFTTSLPKGNVRLTSVKIRECTGLAEMW	436
RESULT	11		
ENTRY		s47297	#type complete
TITLE		sulysin - Streptococcus suis	
ORGANISM		#formal_name Streptococcus suis	
DATE		06-Jan-1995	#sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
ACCESSIONS		s47297	
REFERENCE		s47297	
#authors		Segers, R.	
#submission		submitted to the EMBL Data Library, September 1994	
#accession		s47297	
##status		preliminary	
##molecule_type		DNA	
##residues		1-96	#label SEG
##cross-references		EMBL:236907; NID:g535305; PID:g535306	
SUMMARY		#length 96	#molecular_weight 10471 #checksum 6662
Query Match		6.6%; Score 214; DB 2; Length 96;	
Best Local Similarity		33.7%; Pred. No. 3,28e-15;	
Matches		30; Conservative 23; Mismatches 36; Indels 0; Gaps 0;	
Db	1	KQDNGYFOSLIVYGPQELITNCEGYIDNPATGMLENGFRVYVLRREKXTITDNSADITVI	60
QY	4	KAVNDFILAMNDKRLKLLTHGQESIEFKKGNQLPDEFVYXERKRKSLSTNTSDIXVT	63
Db	61	XAKANISPGALLRAXOVLDDNPLIISI	89
QY	64	ATNDSRLTPGALLVDETLENNPTLAAV	92
RESULT	12		
ENTRY		s03974	#type complete
TITLE		amine oxidase (flavin-containing) (EC 1.4.3.4) A - Bovine	
ALTERNATE_NAMES		monoamine oxidase type A	

```

ORANISM      #journal.name Bos primigenius taurus #common_name cattle
DATE         26-Feb-1990 #sequence_revision 26-Feb-1990 #text_change
ACCESSIONS   S03974
REFERENCE     #authors Powell, J.F.; Hsu, Y.P.; Weyler, W.; Chen, S.; Salach, J.;
              #journal Andrikopoulos, K.; Mallet, J.; Breakefield, X.O.
              #title Biochem. J. (1989) 259:407-413
              The primary structure of bovine monoamine oxidase type A.
              Comparison with peptide sequences of bovine monoamine
              oxidase type B and other flavoenzymes.
              #cross-references MUID:89246344
              #accession S03974
              #status not compared with conceptual translation
              ##molecule_type mRNA
              ##residues 1-527 ##label POW
              ##cross-references GB:X15609; NID:g9523; PID:g524
KEYWORDS      FAD; flavoprotein; mitochondrion; oxidoreductase
FEATURE       15-43
              #region beta-alpha-beta FAD nucleotide-binding fold\
              modified site S (alpha-PAD)-cysteine (Cys) #status
              predicted
SUMMARY       #length 527 #molecular-weight 59800 #checksum 6598

Query Match    3.6%; Score 117; DB 2; Length 527;
Best Local Similarity 19.3%; Pred. No. 1,18e-01;
Matches 29; Conservative 42; Mismatches 73; Indels 6; Gaps 5;

Db 1 MESLQKTDAGCMPDVYVIGGICISGLNAK-LLAHEVNVLVEARENRGGRTTYRNEH 59
||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||::|
QY 250 LETTSXSEVEAEFALLIKGVAPQTEWKQILDNTEVKAVIIGDPPSGARVVTGKVMM 309
||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||::|
Db 60 V-DYDVCGAAYVGPFQNILRLSKQLGETRYKVNVRNLHYKKGTTPRGAFPPWMP 118
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 310 VEDLIQESRFTADHPGLPISTYSF-LRDNVVAATQNSIDIYETVTAIRNGDLLDHS 368
||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||::|
Db 119 IAYL-DYNMLWRFM--DNMKEIPDAPE 145
||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||::|
QY 369 GAYVAQYYITWNELSYDHQGEVLTPRAMD 398

RESULT 13
ENTRY   S59327 #type complete
TITLE   hypothetical protein YLR135w - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES  hypothetical protein Lj140
ORGANISM  #formal.name Saccharomyces cerevisiae
          29-Nov-1995 #sequence_revision 23-Feb-1996 #text_change
          05-Dec-1997
ACCESSIONS  S59327; S64977
REFERENCE     #authors Delius, H.
              #submission submitted to the EMBL Data Library, June 1995
              #description 36.8 kb of S.cerevisiae chromosome XII including ACE2, CKT1,
              PDC5, SLS1, PUT1 and tRNA-asp.
              #accession S59327
              ##molecule_type DNA
              ##residues 1-748 ##label DEL
              ##cross-references EMBL:X91258; NID:g995686; PID:g995701
              #experimental_source strain S288C
REFERENCE     S64967
              #authors Delius, H.; Hebling, U.
              #submission submitted to the Protein Sequence Database, May 1996
              #accession S64977
              ##molecule_type DNA
              ##residues 1-748 ##label DEW
              ##cross-references EMBL:L273307; NID:g1360550; PID:e245572; PID:g1360551;
              MIPS:YLR135w
              #experimental_source strain S288C

GENETICS
#map_position 12R
SUMMARY       #length 748 #molecular-weight 84361 #checksum 5347

```

```

Query Match      3.5%: Score 112: DB 2: Length 748:
Best Local Similarity 18.8%: Pred. No. 4.55e+01:
Matches 33: Conservative 58: Mismatches 71: Indels 14: Gaps 13:

Db 567 SEEEETEDDFOCFIADLIQVLDSSKISTKSDTOPPTTNSIDIIDTSSAASSIASPEKFEIWM 626
      : : : : : | | | | | : : : : : | : | : | : : : : :
Oy 22 THOGS-1ENFXKXEGNOLDPDEFVXKRRKRSLSTNSDLXVATANDSKRYPAL--LVY 78
      : : : : : | | | | | : : : : : | : | : | : : : : :

Db 627 SOSKMKLRQSLKTYGLKPMRTKEIILQIOTASOILSTANPDNKGHGVANF-SK-IEI 684
      : : : : : | | | | | : : : : : | : | : | : : : : :
Oy 79 DELIKENNPILLADRAPMTYSIDL-PGLASSDSFLQVEDPSSSVRGAVNDLLAKHQD 137
      : : : : : | | | | | : : : : : | : | : | : : : : :

Db 685 FDHLIELIENPDLERITYPEEIPILNELLEKLFSAEPFYSQDEITIKEMADYOG 740
      : : : : : | | | | | : : : : : | : | : | : : : : :
Oy 138 YGQVNN-VPARXQY-EKI-T-AH-SMDLKVK-FGSD-FEKTGNSLDI-DENSVHS 185
      : : : : : | | | | | : : : : : | : | : | : : : : :

RESULT 14
ENTRY F64461
TITLE hypothetical protein M01295 - Methanococcus jannaschii
ORGANISM #format_name Methanococcus jannaschii
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 10-Oct-1997

ACCESSIONS F64461
REFERENCE A64300
#authors Bul, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.A.; Fitzgerald, L.M.; Clayton, R.A.; Gocayne, J.D.; Kierlavage, A.R.; Dougherty, B.A.; Tomb, J.F.; Adams, M.D.; Reisch, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glöck, A.; Scott, J.L.; Geoghegan, N.S.M.; Meldrum, J.F.; Furman, J.L.; Nguyen, D.; Ulterback, T.R.; Kelley, J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Rienk, H.P.; Fraser, C.M.; Smith, H.O.; Moese, C.R.; Venter, J.C. Science 1996) 273:1058-1073
#journal Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.
#cross-references M01D:96337999
#accession F64461
#status preliminary: nucleic acid sequence not shown; translation not shown

##molecule_type DNA
##residues 1-218 #label BUL
##cross-references GB:U67570; GB:L77117; NID:g1591929; PID:g1591933; TIGR:M01295; PID:g1511304

GENETICS
#map_position REV1243502-1242846
#start_codon TTG
SUMMARY #length 218 #molecular_weight 24666 #checksum 9893

Query Match      3.3%: Score 107: DB 2: Length 218:
Best Local Similarity 26.3%: Pred. No. 1.68e+00:
Matches 21: Conservative 17: Mismatches 40: Indels 2: Gaps 2:

Db 68 IGVGVELTHIPKSPKMKAKKADGAEIVVH-GETVVEPEVEKTKTYVASISEDVDILAH 126
      : : : : : | | | | | : : : : : | : | : | : : : : :
Oy 153 ITASHMEQLKVKFSSDEPKTGNSLDIDFNSVHSGEKQIQIVNKKQIYIVYSDAKNRPD 212
      : : : : : | | | | | : : : : : | : | : | : : : : :

Db 127 PGFIDKEFLAENMLKENDIFEV 146
      : : : : : | | | | | : : : : : | : | : | : : : : :
Oy 213 V-FQDTYVVEDLKKORGISAE 231
      : : : : : | | | | | : : : : : | : | : | : : : : :

RESULT 15
ENTRY S60947
TITLE hypothetical protein YOR220w - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES hypothetical protein OS050: hypothetical protein YOR50-10
ORGANISM #format_name Saccharomyces cerevisiae
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 05-Jun-1998
ACCESSIONS S60947; S67113; S71722

```

```

REFERENCE      S60938
#authors      Gallison, F.; Dujon, B.
#submission   submitted to the EMBL Data Library, October 1995
#description   Sequence and analysis of a 33 kb fragment from the right arm
               of chromosome XV of the yeast Saccharomyces cerevisiae.
#accession    S60947
               ##molecule-type DNA
               ##residues      1-265 ##label GAL
REFERENCE      S67104
#cross-references EMBL:X92441; NID:g1050762; PID:g1050772
#authors      Boyer, J.; Fairhead, C.; Gallion, L.; Gallison, F.; Michaux,
               G.; Thierry, A.; Dujon, B.
#submission   submitted to the Protein Sequence Database, July 1996
#accession    S67113
               ##molecule-type DNA
               ##residues      1-265 ##label BOY
               ##cross-references EMBL:Z75128; NID:g1420509; PID:e252085; PID:g1420510
               MFS:YOR820W
REFERENCE      S71713
#experimental-source strain S288C
#authors      Gallison, F.; Dujon, B.
#journal      Yeast (1996) 12:877-885
#title        Sequence and analysis of a 33 kb fragment from the right arm
               of chromosome XV of the yeast Saccharomyces cerevisiae.
#cross-references MUID:96437977
#accession    S71722
               ##status      nucleic acid sequence not shown; translation not shown
               ##molecule-type DNA
               ##residues      1-265 ##label GAW
               ##cross-references EMBL:X92441; NID:g1050762; PID:g1050772
               ##note        the nucleotide sequence was submitted to the EMBL Data
               Library, October 1995

GENETICS
#map-position 15R
SUMMARY      #length 265 #molecular-weight 29255 #checksum 6028

Query Match      3.3%; Score 106; DB 2; Length 265;
Best Local Similarity 26.7%; Pred. NO. 2.17e+00;
Matches 28; Conservative 26; Mismatches 44; Indels 7; Gaps 7;

Db 85 PAKLFVETSLNNO-HRR-SRSTDDAVSLQDNNLALLEDRKKPFLSINTDQGVTVDS 142
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 12 AMN-YDKKKLLTHGSGEINERFKREGQLPDE-FVYXF-RKKRSLSTNNSDXIVATINDS 68
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 143 SLNKGSSLSDFKSSLSFPTMLKISTDSKFSYQEDLPKLSRSS 187
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 69 RLY-PGALLVVDFTLKENNPTL-LAVDRAPMYTSIDLPGLASSDS 111

Search completed: Fri Sep 3 13:11:27 1999
Job time : 48 secs.

```

Search completed: Fri Sep 3 13:11:27 1999
Job time : 48 secs.



Matches 460; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Db 1 ANKAVNDPILAMNDKRLTHGSEIENRFINEGNLPDEFVYIERKKRSLSTNTSDIS 60
 QY 2 ANKAVNDPILAMNDKRLTHGSEIENRFINEGNLPDEFVYIERKKRSLSTNTSDIX 61
 Db 61 VTANDSRIVPGALVVDPELLENNPLLAVDAPMTYSIDLGLASSDSFLQVDEPSSNS 120
 QY 62 VTANDSRIVPGALVVDPELLENNPLLAVDAPMTYSIDLGLASSDSFLQVDEPSSNS 121
 Db 121 SVRGAVDLLAKMHQDYGAVNNPAPAKQYKETAHSMEOQKVFSGDFEKTGSLDIDFN 180
 QY 122 SVRGAVDLLAKMHQDYGAVNNPAPAKQYKETAHSMEOQKVFSGDFEKTGSLDIDFN 181
 Db 181 SVHSGEQIOIVNFQOITYYVSVDVAKNPGDVFQDITVEDLKQRGISAEPLVISSVA 240
 QY 182 SVHSGEQIOIVNFQOITYYVSVDVAKNPGDVFQDITVEDLKQRGISAEPLVISSVA 241
 Db 241 YGRQVYLKLETTSSDVEAAFEALIKGVKAPQTEWKQILDNTEVKAIVLLGGDPSSGAR 300
 QY 242 YGRQVYLKLETTSSDVEAAFEALIKGVKAPQTEWKQILDNTEVKAIVLLGGDPSSGAR 301
 Db 301 VVTKGVMDVEDLIOESRFTADHGLPISTTSFLRDNVATFONSTDYETKVTAYRNG 360
 QY 302 VVTKGVMDVEDLIOESRFTADHGLPISTTSFLRDNVATFONSTDYETKVTAYRNG 361
 Db 361 DLLIDHSGAVYAOYITWDELSDYHOGKEVLTFRAMDRNGODLTAHFTTISPLKGNVRL 420
 QY 362 DLLIDHSGAVYAOYITWDELSDYHOGKEVLTFRAMDRNGODLTAHFTTISPLKGNVRL 421
 Db 421 SVKIRECTGLAMWMTVYEKTDLPYRKRTISMGTLTPQVEDKEND 470
 QY 422 SVKIRECTGLAMWMTVYEKTDLPYRKRTISMGTLTPQVEDKEND 471

RESULT 2
 ID TACY CLOPE STANDARD; PRT; 500 AA.
 AC P19995;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE PERRINGOLYSIN O PRECURSOR (THETA-TOXIN) (THIOL-ACTIVATED CYTOLYSIN).
 GN PRC OR PROR OR PROA.
 OS CLOSTRIDIUM PERRINGENS.
 OC PLASMD PR1B.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;
 CC CLOSTRIDIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 9109951.
 RA SHIMIZU T., OKABE A., MINAMI J., HAYASHI H.;
 RT "An upstream regulatory sequence stimulates expression of the
 RT perrinngolysin O gene of Clostridium perringens.";
 RL INFECT. IMMUN. 59:137-142(1991).
 RN [2]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 29-44.
 RX STRAIN-ATCC 13124;
 RX MEDLINE; 89032623.
 RA TWETEN R.K.;
 RT "Nucleotide sequence of the gene for perrinngolysin O (theta-toxin)
 RT from Clostridium perringens: significant homology with the genes for
 RT streptolysin O and pneumolysin.";
 RL INFECT. IMMUN. 56:3235-3240(1988).
 RN [3]
 RP SEQUENCE OF 29-45 AND 305-312.
 RX STRAIN-PB6K;
 RX MEDLINE; 87076517.
 RA CHINO-IWASHITA Y., IWAMOTO M., MITSUI K., KAWASAKI H., ANDO S.;
 RT "Cold-labile hemolysin produced by limited proteolysis of theta-toxin
 RT from Clostridium perringens.";
 RL BIOCHEMISTRY 25:6048-6053(1986).
 RN [4]
 RP SEQUENCE OF 492-500 FROM N.A.

RC STRAIN-NCTC 8237;
 RX MEDLINE; 96123363.
 RA SHIMIZU T., KOBAYASHI T., BA-THEIN W., OHTANI K., HAYASHI H.;
 RT "Sequence analysis of flanking regions of the pfoA gene of
 RT Clostridium perringens: beta-galactosidase gene (pbg) is located in
 RT the 3'-flanking region.";
 RL MICROBIOL. IMMUNOL. 39:677-686(1995).
 RN [5]
 RP IMPORTANCE OF THIOL-GROUP IN CHOLESTEROL BINDING.
 RX MEDLINE; 88004463.
 RA IWAMOTO M., OHNO-IWASHITA Y., ANDO S.;
 RT "Role of the essential thiol group in the thiol-activated cytolysin
 RT from Clostridium perringens.";
 RL EUR. J. BIOCHEM. 167:425-430(1987).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE; 97113505.
 RA FEIL S.C., ROSSJOHN J., ROHDE K., TWETEN R.K., PARKER M.W.;
 RT "Crystallization and preliminary x-ray analysis of a thiol-activated
 RT cytolysin.";
 RL FEBS LETT. 397:290-292(1996).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE; 97325744.
 RA ROSSJOHN J., FEIL S.C., MCKINSTRY W.J., TWETEN R.K., PARKER M.W.;
 RT "Structure of a cholesterol-binding, thiol-activated cytolysin and a
 RT model of its membrane form.";
 RL CELL 89:685-692(1997).
 CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
 CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 CC EUKARYOTIC CELL MEMBRANES.
 CC -1- SUBUNIT: FORMS OLIGOMERS IN THE HOST MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----

DR EMBL; M81080; G144886; -;
 DR EMBL; M36704; G144884; -;
 DR EMBL; D49537; G1502275; -;
 DR PIR; B43577; B43577.
 DR PDB; 1PFO; 05-AUG-98.
 DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
 DR PFAM; PF01289; THIOL_Cytolysin; 1.
 KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMD; 3D-STRUCTURE.
 FT SIGNAL 1 28
 FT CHAIN 29 500 PERRINGOLYSIN O.
 FT ACT SITE 459 459 BINDING TO CHOLESTEROL.
 FT CONFLICT 30 30 D -> K (IN REF. 3).
 FT CONFLICT 34 34 K -> I (IN REF. 3).
 FT CONFLICT 126 128 RKP -> EA (IN REF. 2).
 SQ SEQUENCE 500 AA; 55799 MW; 17FC32BC CRC32;

Query Match 49.5%; Score 1600; DB 1; Length 500;
 Best Local Similarity 46.5%; Pred. No. 0.00e+00;
 Matches 214; Conservative 105; Mismatches 140; Indels 1; Gaps 1;

Db 35 NOSIDGSSLSYNEVLAASNGDKIESFVEKGGKAKNFIYEROKRSLTSPVDISI 94
 QY 3 NKAVDNFIAMNDKRLTHGSEIENRFINEGNLPDEFVYIERKKRSLSTNTSDIX 62
 Db 95 IDSVDRTYPPGALADAFVFNRPITIMVAKKININIDLPGLKGENS-1KYDDPTYGK 153
 QY 63 TATNSRRLYPGALLVVDDELTXENNPPLLAVDAPMTYSIDLGLASSDSFLQVDEPSSNS 122
 Db 154 VSGAIDELVSKWNEKYSSTHTLPARTOYSESMVYSKSOISSALVWNAKYLENSIGVDFNA 213

```

CC Modfile and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL/ X15127; G44107; -;
DR EMBL/ M24199; G149653; -;
DR EMBL/ X60035; G44112; -;
DR EMBL/ U25452; G687870; -;
DR EMBL/ U25443; G687016; -;
DR EMBL/ U25446; G687028; -;
DR EMBL/ U25449; G687864; ALT_INIT.
DR PIR/ S05306; S05306;
DR PIR/ A43505; A43505;
DR PROSITE/ PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM/ PF01289; Thiol_cytolysin; 1.
DR HSSP/ P19995; 1PFO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 529 LISTERIOLYSIN
FT ACT_SITE 484 484 BINDING TO CHOLESTEROL (BY SIMILARITY).
FT VARIANT 35 35 S->L (IN STRAINS F2365, F4233, F5782,
FT F6789 AND 12067).
FT VARIANT 39 39 M->I (IN STRAIN F2365).
FT VARIANT 438 438 V->I (IN STRAINS F2365, F4233, F5782,
FT F6789 AND 12067).
FT VARIANT 523 523 K->S (IN STRAINS F2365, F4233, F5782,
FT F6789 AND 12067).
SQ SEQUENCE 529 AA; 58688 MM; 269EA737 CRC32;
Query Match 47.7%; Score 1543; DB 1; Length 529;
Best Local Similarity 42.9%; Pred. No. 7, 51e-301;
Matches 200; Conservative 117; Mismatches 148; Indels 1; Gaps 1;
Db 61 IDKYQGLDYNNKNNLVYHGDAVTNVPKRGKGDENEYIYVEKKKKSINONNADIOYVNA 120
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 6 VNDFTLANNYDKKILLTHQGESIENRFKKEGQNPDEVEYAKKKRSLSTNSDIXVAT 65
121 ISSLTYPGALVANKSELVENOPVLVPRKRLTSLISIDLPGMTNODNKIVYKNAKTSNVNN 180
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 66 NDSRLYPGALLVVDETLKENNPILLAVRAPMTYSIDPLGLASSDSFLVEDPSSNSVVG 125
181 AVNTIYVERKNEKYAOAYRNVNSAKIDYDDEMAVSEQLAKFGTAKAVANSLNVFGAIS 240
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 126 AVNDLALWHODYGOVN-NVPARKQYEKITAHSMQLVYKFGSDDEKGTSLDIDFNSVH 184
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 241 EKKMOEVIYSFKOITYYNNVNEPTPRKRFEGKAVYKEDOLALGVNAENPPAYISSVAYGR 300
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 165 SGEKQIQIVNKKQIITYYSVDAYKNRPGDVFODTYVEDLKORGISIAERPLYIISXVAIGR 244
301 OYVTLKSTNSHSTKVAAFDAAVSGSKSVSGVELNIIKNSSEKAVIYVGSAGKDEQVQIT 360
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 245 OYVTLKSTNSHSTKVAAFDAAVSGSKSVSGVELNIIKNSSEKAVIYVGSAGKDEQVQIT 360
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 361 GNLGLDRDLIKKGAFFNKPETPGVPLAITTNPLKDNELAVIKNNSIETSKAAYTDGKIN 420
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 305 GKYDVEDELLDGGSFPTADHPGLPSTIYTSFLRDVVVATFONSTYVETTKYVARNQDL 364
421 IDHSGGYAAQFIISMDEVNDYDEGNEIYOHKRWKSNNSKLAHPFSSIYLPGNANINVY 480
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 365 LDHSGGYAAQFIITNNELSYDHQGEVLLPRAKMDRNGODLAHFTSTIDPLKGNVHNSLVK 424
481 AKECTGLAMEMWRTYIDDRNLPLVYNNRNISIGTLLPYKTSKVN 526
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 425 IREGTGLAMEMWRTYVEKTDPLVVKRTISINGTLLPYQVEDKVEN 470
RESULT 4
ID TACY_LISTIV STANDARD; PRT; 528 AA.
AC P31831;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE IVANOIYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).

```


[illegible]

```

OY      130 LLAWHODVGVNNVPARXQYEKRTIASHMEQLKFCSDFEKTNSLIDIFNSVHSEKQ 189
Db      223 VMVAAYKOIFTVSAGLPNNRSDFLDPSYTPFAELARGVSNAPRLMVAANAVERITTVK 282
OY      190 IOIVMMKOYYTVDADKANPDGFODITYVEDLKORGISAEPLVYISXYAAGRQYLK 249
Db      283 LETTSKSNDVOCTAFELLNPNFSIOASGGOKRIENSSFTAVLGDDAQTHNQVTKDENV 342
OY      250 LETTSKSXEVEALPEALKGVKAPQTOMKOLINTIEVKAVILGDDSNGARVYTGAADM 309
Db      343 IOSVIKDAOSKSKNPAPISTYSVFLEKDNSIAAVHNNTXIETKTEYSKGIRKLDSHG 402
OY      310 VEDLIQESREFTRADHPGLPISYTTFSLDNVANAFONSTDYERKVTAYRNGLDLLDSHG 369
Db      403 AYVAOFEVYMDEFSDADGOELVTRKSMDGWNRBRSASFSEIPLPRANKIRLFARPCT 462
OY      370 AYVAOYTYTMVELLSHDHOGKEVLTPRKADRNGODTLTAFTSIPLGKNVM,LVKIRECT 429
Db      463 GLAMEMRWTVDEVNVPLASDLINSGTTLYP 495
OY      430 GLAMEWRVIEKLDPLVRKRRTISMGTTTTYP 462

RESULT       7
ID           TACY_STRSEQ    STANDARD;          PRF;     574 AA.
AC           Q54114;
DT           15-JUL-1998 (REL. 36, CREATED)
Dt           15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE           15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
GN           STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
OS           SLO.
OC           STREPTOCOCCUS EQUISTIMILIS.
OC           BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE.
CN           STREPTOCOCCUS.
RN           [1]
RP           SEQUENCE FROM N.A.
RC           STRAIN-SIMD-1.
RX           MEDLINE: 95102113.
RA           OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
            IGARASHI H., YUTSUJO T.;
RT           "Cloning and sequencing the streptolysin O genes of group C and group
            G streptococci."
RL           DNA SEQ. 4:325-328(1994).
CC           -I- FUNCTION: SUPPHDRIL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
            CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
            CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
            EUKARYOTIC CELL MEMBRANES.
CC           -I- SIMILIARTY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
Cc         -----
Cc         This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc         between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc         the European Bioinformatics Institute. There are no restrictions on lists
Cc         use by non-profit institutions as long as its content is in no way
Cc         modified and this statement is not removed. Usage by and for commercial
Cc         entities requires a license agreement (See http://www.isb-slb.ch/announce/
Cc         or send an email to licenselsb-slb.ch).
Cc         -----
Cc         EMBL: D16824; G498301; -.
Dr         PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
Dr         PEAR: PF01289; ThiOL_CytoLySins; 1.
Dr         HSP: p19995; lpro.
KW         TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT         CHAIN             1              36               BY SIMILARITY.
FT         ACT SITE        37              574             STREPTOLYSIN O.
FT         SITE            533             BINDING OF CHOLESTEROL (BY SIMILARITY).
SQ         SEQUENCE        574 AA;   63991 MW;   9BBF31B4 CRC32;

Query Match                43.6%; Score 1409; DB 1; Length 574;
Best Local Similarity 41.6%; Pred.No. 4.24e-271;
Matches 190; Conservative 106; Mismatches 160; Indels 1; Gaps 1;

```

```

QY 6 VNDPILAMNDKKLLTHOGESIENFRKXEGNQLPDEFVYXERKRSLSTNTSDIXVAT 65
DB 172 VYDRTPAALOLANKGFENKPDVAVTKRNPQKHIDLPDGMCKAT-VEVNDPTANST 230
QY 66 NDRSLYPGALLVYDELXENNPTLLAVDAPMTYSIDLPGGLSSSFQEDPSGSRVVG 125
DB 231 AIDNLVNMHDNYSNGNTLPARTQYTESNYSKSOIEALNYSKILDTGLGIDFKSISK 290
QY 126 AVNDLLAKHODYGVYNNVAPARQYKITAHSMEQLKVFSGDFEKTGNSLIDFNSVHS 185
DB 291 GEKVMIAAOKIEFYVSANLPNNPADVDKSVTLKEIQRKVSNEAPPLVSNVAGRT 350
QY 186 GEKOIOIVXKKOITYTVSDAVKNPQDVFDVTEDEKORGISAEPLVYISXVAVGRQ 245
DB 351 VEVKLETSSKSDNDEAFAALKGTVDKNGKSDILENSSFTAVVLGDAEHNKVVTK 410
QY 246 VYLKLETTSSXSEVEAFALIKGVKVAPOTEWKQILDNTEKAVIILGDSGSRVVG 305
DB 411 DEDVIRNVYKDNATSRKRPAPVITSYFELKNNKIAGVNNREVEVETTSYTGKINTL 470
QY 306 KYDMVEDLIQESRFTADHPGLPISTYSFLRDVNAVFQNSTDYVETKVAAYRNGDLL 365
DB 471 SHOGAYVAQYEILMDEINVDCKEVIKRRMDNNWYSKTSFSPVITPLGANSRIRIMA 530
QY 366 DHSAGVVAQYITTWMLSDYDHOGKEVLPKADRNGQDLTAHFTTSLPKGNVNLVYKI 425
DB 531 RECTGLAMEMWRKVIDERDVKLSEINYNISGSLSP 567
QY 426 RECTGLAMEMWRVYKEDLPLVRKRRTISIMGTILYP 462

RESULT 8
ID TACY_STRPY STANDARD: PRT: 571 AA.
AC P21131.
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS PYOGENES.
OC PLASMID PMK157.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
RN STREPTOCOCCUS.
RP [1]
RA SEQUENCE FROM N.A.
RA MEDLINE: 88057628.
RA KEROE M.A., MILLER L., WALKER J.A., BOULNOIS G.J.;
R1 "Nucleotide sequence of the streptolysin O (SLO) gene: structural
homologies between SLO and other membrane-damaging, thiol-activated
toxins."
RT INFECT. IMMUN. 55:3228-3232(1987).
CC -!- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -!- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: M18638; G153811; .
DB PIR: A43507; A43507.
DB PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
DB PRAM: PFO1289; Thiol_cytolysin; 1.
DB HSP: P19995; 1PFO.
KW TOXIN, HEMOLYSIS, CYTOLYSIS, SIGNAL, PLASMID.
FT SIGNAL 1 33
CHAIN 34 571 STREPTOLYSIN O.

```

```

FT ACT SITE 530 530 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 571 AA: 63638 MW: 3312454 CRC32:
Query Match 43.5%; Score 1408; DB 1; Length 571;
Best Local Similarity 41.8%; Pred. No. 7,078-271;
Matches 191; Conservative 104; Mismatches 161; Indels 1; Gaps 1;

DB 109 INKISLANTNELEVALAKGETIENVEPEGVKADKFLVIERKKNNINTPVDISIDS 168
QY 6 VNDPILAMNDKKLLTHOGESIENFRKXEGNQLPDEFVYXERKRSLSTNTSDIXVAT 65
DB 169 VYDRTPAALOLANKGFENKPDVAVTKRNPQKHIDLPDGMCKAT-VEVNDPTANST 227
QY 66 NDRSLYPGALLVYDELXENNPTLLAVDAPMTYSIDLPGGLSSSFQEDPSGSRVVG 125
DB 228 AIDNLVNMHDNYSNGNTLPARTQYTESNYSKSOIEALNYSKILDTGLGIDFKSISK 287
QY 126 AVNDLLAKHODYGVYNNVAPARQYKITAHSMEQLKVFSGDFEKTGNSLIDFNSVHS 185
DB 288 GEKVMIAAOKIEFYVSANLPNNPADVDKSVTLKEIQRKVSNEAPPLVSNVAGRT 347
QY 186 GEKOIOIVXKKOITYTVSDAVKNPQDVFDVTEDEKORGISAEPLVYISXVAVGRQ 245
DB 348 VEVKLETSSKSDNDEAFAALKGTVDKNGKSDILENSSFTAVVLGDAEHNKVVTK 407
QY 246 VYLKLETTSSXSEVEAFALIKGVKVAPOTEWKQILDNTEKAVIILGDSGSRVVG 305
DB 408 DEDVIRNVYKDNATSRKRPAPVITSYFELKNNKIAGVNNREVEVETTSYTGKINTL 467
QY 306 KYDMVEDLIQESRFTADHPGLPISTYSFLRDVNAVFQNSTDYVETKVAAYRNGDLL 365
DB 468 SHOGAYVAQYEILMDEINVDCKEVIKRRMDNNWYSKTSFSPVITPLGANSRIRIMA 527
QY 366 DHSAGVVAQYITTWMLSDYDHOGKEVLPKADRNGQDLTAHFTTSLPKGNVNLVYKI 425
DB 528 RECTGLAMEMWRKVIDERDVKLSEINYNISGSLSP 564
QY 426 RECTGLAMEMWRVYKEDLPLVRKRRTISIMGTILYP 462

RESULT 9
ID TACY_STRCB STANDARD: PRT: 574 AA.
AC Q53957.
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS CANIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
RN STREPTOCOCCUS.
RP [1]
RA SEQUENCE FROM N.A.
RA MEDLINE: 95102113.
RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
R1 "Cloning and sequencing the streptolysin O genes of group C and group
RT G streptococci."
CC -!- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -!- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```


CC -1- FUNCTION: CATALYZES THE OXIDATIVE DEMINATION OF BIOGENIC AND
 CC XENOBOTIC AMINES AND HAS IMPORTANT FUNCTIONS IN THE METABOLISM OF
 CC NEURACTIVE AND VASOACTIVE AMINES IN THE CENTRAL NERVOUS SYSTEM
 CC AND PERIPHERAL TISSUES. MAO-A PREFERENTIALLY OXIDIZES BIOGENIC
 CC AMINES SUCH AS 5-HYDROXYTRYPTAMINE (5-HT), NORPINEPHRINE AND
 CC EPINEPHRINE.
 CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
 CC H(2)O(2).
 CC -1- COFACTOR: FAD.
 CC -1- SUBUNIT: MONOMER, HOMO- OR HETERODIMER (CONTAINING TWO SUBUNITS OF
 CC SIMILAR SIZE). EACH SUBUNIT CONTAINS A COVALENTLY BOUND FLAVIN.
 CC ENZYMATICALLY ACTIVE AS MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL OUTER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X15609; G524; -
 CC DR EMBL; X15609; G525; ALT_INIT.
 CC DR PIR; S03974; S03974.
 CC KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; TRANSMEMBRANE; MITOCHONDRION;
 CC NEUROTRANSMITTER DEGRADATION.
 CC FT NP_BIND 14 70 FAD (ADP PART) (POTENTIAL).
 CC FT BINDING 406 406 FAD (BY SIMILARITY).
 CC FT TRANSMEM 498 518 POTENTIAL.
 CC SQ SEQUENCE 527 AA; 59800 MW; 724E4396 CRC32;
 CC
 CC Query Match 3.68; Score 117; DB 1; Length 527;
 CC Best Local Similarity 19.38; Pred. No. 2,05e-02;
 CC Matches 29; Conservative 42; Mismatches 73; Indels 6; Gaps 5;
 CC
 CC Db 1 MESLQTS DAGQMFVYVVGIGISGLSAK-LLAEHEVVLLEARENGGTYVREH 59
 CC QY 250 LETTSXSEVEAFELINGVAVAPOTEMKQILDNTEVAAILGDPSSGARVVGKVD 309
 CC Db 60 V-DYDVGVGAYVPTONRLRLSKQLGLETYYVNVNERLVHYKGTYPFGCAFPPVWNP 118
 CC QY 310 VEDLIDGSGRTADHGLPISTYTSF-LRDNVVATFQNSTDYEVKRYAAYRNGDLLDHS 368
 CC Db 119 IAYL-DYNNLMRTM--DNMGKEIPADAPWE 145
 CC QY 369 GAVVAQYITWNLSTYDHQKEVLTPKAWD 398
 CC
 CC RESULT 12
 CC ID MPEE_YEAST STANDARD: PRT: 462 AA.
 CC AC P10507;
 CC DT 01-JUL-1988 (REL. 11, CREATED)
 CC DT 01-JUL-1988 (REL. 11, LAST SEQUENCE UPDATE)
 CC DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 CC DE MITOCHONDRIAL PROCESSING PEPTIDASE BETA SUBUNIT PRECURSOR
 CC (EC 3.4.24.64) (BETA-MP) (PEP)
 CC GN MASI OR MFI1 OR YLR163C OR L9632.10.
 CC OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
 CC OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACHAROMYCETALES;
 CC OC SACHAROMYCETACEAE; SACHAROMYCES.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE: 88312592.
 CC RA WITTE C., JENSEN R.E., YAFFE M.P., SCHATZ G.;
 CC RT "MAS1, a gene essential for yeast mitochondrial assembly, encodes a
 CC subunit of the mitochondrial processing protease.";
 CC RL EMBL J. 7:1439-1447(1988).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-SE288C / AB972;
 CC RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,

RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J.,
 RA KUCABA T., HALLSWORTH K., HAWKINS J., HILLIER L., JIER M.,
 RA JOHNSON D., JOHNSTON L., LANGSTON Y., LATREILLE P., LE T.,
 RA MADDIS E., MENESSES S., MILLER N., NHAN M., PAULEY A., PELUSO D.,
 RA RIKKEN L., RILES L., TAICH A., TREVASKIS E., VIGNATI D.,
 RA WILCOX L., WOHLDMAN P., VAUDIN M., WILSON R., WATERSTON R.;
 RA SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC [3]
 CC RN SEQUENCE OF 21-32.
 CC RP MEDLINE: 91177897.
 CC RA YANG M., GELI V., OPLIGER W., SUDA K., JAMES P., SCHATZ G.;
 CC RT "The MAS-encoded processing protease of yeast mitochondria.
 CC Interaction of the purified enzyme with signal peptides and a
 CC purified precursor protein.";
 CC RL J. BIOL. CHEM. 266:6416-6423(1991).
 CC -1- FUNCTION: THE MITOCHONDRIAL PROCESSING PROTEASE CLEAVES
 CC PRESEQUENCES FROM MITOCHONDRIAL PROTEIN PRECURSORS.
 CC -1- CATALYTIC ACTIVITY: RELEASE OF N-TERMINAL TRANSIT PEPTIDES FROM
 CC PRECURSOR PROTEINS IMPORTED INTO THE MITOCHONDRION, TYPICALLY WITH
 CC ARG IN POSITION P2.
 CC -1- COFACTOR: REQUIRES DIVALENT CATIONS FOR ACTIVITY.
 CC -1- SUBUNIT: HETERODIMER OF ALPHA AND BETA SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
 CC INSULINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X07649; G3887; -
 CC DR EMBL; U51921; G1234852; -
 CC DR PIR; S00552; S00552.
 CC DR PIR; A38734; A38734.
 CC DR SGD; L0001026; MASI.
 CC DR PROSITE; PS00143; INSULINASE; 1.
 CC DR PFAM; PF00675; Insulinase; 1.
 CC KW HYDROLASE; METALLOPROTEASE; ZINC; MITOCHONDRION; TRANSIT PEPTIDE.
 CC FT TRANSIT 1 20 MITOCHONDRION.
 CC FT CHAIN 21 462 MITOCHONDRIAL PROCESSING PROTEASE
 CC FT METAL 70 70 BETA SUBUNIT.
 CC FT ACT SITE 73 73 ZINC (BY SIMILARITY).
 CC FT METAL 74 74 ZINC (BY SIMILARITY).
 CC FT METAL 150 150 ZINC (BY SIMILARITY).
 CC SQ SEQUENCE 462 AA; 51083 MW; 438D1777 CRC32;
 CC
 CC Query Match 3.3%; Score 107; DB 1; Length 462;
 CC Best Local Similarity 22.3%; Pred. No. 4.15e-01;
 CC Matches 33; Conservative 44; Mismatches 61; Indels 10; Gaps 9;
 CC
 CC Db 68 TAFLEHLAFK-GTQ-NRSQGIETLEIENI-GSHLNAVTSRENTYVYAKSLQEDIPKAVD 124
 CC QY 154 TASHMKQLVKFKGSDPEKGNLSDIDFNHSGEKOQIVNKKOYLYIVSV-DAKKNGD 212
 CC Db 125 IISDLITKVLNLSAIERRDYIIESEVYDKMYDEVFHDHLEHTTYDQPLGRIIIGPT 184
 CC QY 213 VFQDIYVTEDLKRGISARPLVYI-Sx-VA--YGRQYVYLRK-LFTTSXSEVEAFELI 267
 CC Db 185 KNKISITRDLKDYIFKNYKGDRAVLG 212
 CC QY 268 KGKVAAPQTEWKO-IIDNTEVAAILG 294
 CC
 CC RESULT 13
 CC ID NOPS_YEAST STANDARD: PRT: 511 AA.
 CC AC Q12499;
 CC DT 01-NOV-1997 (REL. 35, CREATED)
 CC DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

[illegible]

RESULT	2	PRELIMINARY:	PRT:	497 AA.
AC	085102;			
DT	01-NOV-1998 (TREMBL)	08	CREATED	
DT	01-NOV-1998 (TREMBL)	08	LAST SEQUENCE UPDATE	
DT	01-NOV-1998 (TREMBL)	08	LAST ANNOTATION UPDATE	
DE	HEMOXYLIN.			
OS	STREPTOCOCCUS SUTS.			
OC	BACTERIA: FIRMICUTES: BACILLUS/CLOSTRIDIUM GROUP: STREPTOCOCCACEAE			
OC	STREPTOCOCCUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=1933;			
RT	OKUMABABA O.;			
RT	"streplococcus suis type 2 hemolysin (sullysin) gene.";			
RL	SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBD DATA BANKS.			
DR	EMBL: AF043556; G3309561; ..			
SO	SEQUENCE	497 AA;	54778 MW;	36DAB184 CRC32;

Query Match	54.7%;	Score 1770;	DB 2;	Length 497;
Best Local Similarity	49.8%;	Pred. No. 0.00e+00;		
Matches 233;	Conservative 108;	Mismatches 125;	Indels 2;	Gaps 2

Db	29	SKODINOFQSLTTPPELILNEGEY IONPATGMLENGFEVYLRRKKNITNSADIA	88
Oy	2	AKKANVDIFLANNYDKKLLHOGESINRFXKEGNQJDEFEVYXERRKRSLTNTSDIX	61
Db	89	VIDAKAANIYFPALLRAQONLIDNNPFLTISARGDLTSLPLDGLANDSHITVNSGTRS	148
Oy	62	YVATNDSRLYFPELLVDEITLKENNPFTLLAVDRAPMYISIDPLGLASDSDLOVEDPSNS	121
Db	149	TYRTGVNNLLSKMNTTYGCEYGTNOALLOYDETMAYSQSOLTKTEGTSFEKFIAPVLDINF	208
Oy	122	SVRGVANDLLAKMHODY-GOVNPNPARQOYKRIHAHSEOLKVFSGDEFEKTSNDIDF	180
Db	209	DAVNSGEKOVQIVNKOIYTYTVSDEPSEPSKLAEGTWTEDLKRNOITDEVPVYSSY	268
Oy	181	NSVHSGEKOIQIVNKKOIIYTVSDAVKANPBDYQDIYVYEDLKORGISERPLVYISXV	240
Db	269	SYGRSMFKLETSRSTFOVQAAPFAAIKGVDISGNAEYODILKNTSPSAYIFGDAGSAA	328
Oy	241	AYGRQVYIKLETTSSXSXEVEAEPFALLIKGVYAQOTEMKOLIDNTEVKAIVILGDPSSGA	300
Db	329	TVVSGNIETLTKITEEGRGARYKLBGVDISSTFVVDKNRPAOLLSSEIETFTSYHNS	368
Oy	301	RYVICKGVMDWEDLQOEGSRFTADHPGLISTTYSFLDNVAVFQNSTDVEIKRYAYNR	360
Db	389	SALILDHSGAVYAKYNTITWEEVSYNAGEEYWEKAMDKNGVNLISHMSETIOPGNAR	448
Oy	361	GDLLIDHSGAVYAOYIYTWNELSDYHOGKEVYLAPKAWDRNGODTLAFTSIPLKNOYNR	420
Db	449	LAVNIIOECTGLAMEWMTYVDK-DLPIVQGRKTIWGTITTYIPQADVEY	495
Oy	421	LSVATIREOTGLAMEWMTYVEKTDPLPKRRTJISIMGTITLYPOVEKRY	468

RESULT	3		
ID	031241	PRELIMINARY;	PRT; 534 AA.
AC	031241;		

DI 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PYOLYSIN.
GN PYO.
OS ARCANOBACTERIUM PYOGENES.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIADAE;
OC ACTINOMYCETALES; ACTINOMYCINAEAE; ACTINOMYCETACEAE; ARCANOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN-BBR1
RX MEDLINE: 97464437.
RA BILLINGTON S.T., JOST B.H., CUEVAS W.A., BRIGHT K.R., SONGER J.G.;
RT "The Arcanobacterium (Actinomycetes) pyogenes hemolysin, pyolysin, is a
RT novel member of the thiol-activated cytolysin family".
RL J. BACTERIOL. 179:6100-6106(1997).
DR EMBL: U84782; G252800; -
DR PFAM: PF01289; Thiol_cytolysin; 1.
SQ SEQUENCE 534 AA; 57873 MW; B0C14CA8 CRC32;

Query Match	40.3%;	Score 1302;	DB 2;	Length 534;
Best Local Similarity	40.0%;	Pred. No. 3,48e-24;		
Matches	189;	Conservative 121;	Mismatches 154;	Indels 8;
			Gaps	6

```

Db      63 ETGVDKYIRSLKYPDPSGVLAVKGESEIENVPTK-DQLKGTITYVTKRHKSKFNNLRSDIS 121
      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY      3 NKAIVDFILAMNYDKKKLLTHOGESEIENRFKKEGNQLPDE-FVYXERKKRSISTNTSGDIx 61

```

```

Db      122 AFDANNAHYPGALVLANKDLAKGSPISIGIARPQTVSYVDLPGLVGDKNKVVINNPTKS 181
      :|:::|||||:::  |  ::::|||  |::|:::  :  :::::|
QY      62 VTATNDSRLPGALLVDETLXENNPILLAYDRAPMTYSIDLPGLAASSDSFLQVEDEPSN 121

```

```

D6      182  SYNGGLNGLLDGMQRNSKYDHAAKSYDEIWTSKRPLEAKLGIFEKYSAKLNDVD 241
      11  : : 11 11 : : : : 11 : : 11 11 : : 11 : : 11 : : 11 : :
QY      122  SYGCAVNDLAKWHDYGYQVANNVPARKQYEKITAHSMEQLKVKGSDFEKGTGNSLDIDEN 181

```

```
D6 242 AIHKREROVALASFQOIYYTASVDTPISPHSVFGPNVTAODLKBKRGVNKNKPGLGYSSVS 30D  
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Q7 182 SVHSGEKQLQIVNKKQOIIYTVSVDANKNPBGDVFOQDTIVEDLKORGISAEPRPLVYTSXVA 241I
```

Db 302 YGRQIFVKLETTSTSDNDVQAQAFSGLEKRAKFGUNSTEEFRAKADILNKTRATVYAVGSGAR 361
 ||||:||||| : :||| :||: : : : :||:|
 QY 242 YGRQVYLKLETTSTSXSEVEAFAEALIKG--VKVAPQ--TEWKQIILDNTVEKAVILGSDPS 297

Db 362 GGATVATGIDALKKIIKESTYSTVPAVPVSYAVNFKDNQIAA VRSSGDIETTATT 421
: | | | | : | | | : | | | : | | | : | | | :
QY 298 SGARVVTGKVDNVEDLIQESRFTADHPGLPISTYTSLRDNVATFQNSTDYVEIKVTA 357

Dd 422 YXSGEITRRHGGGVAAKFRKLWDEISYDPQGEIRTPETWSGNNAARTGFEETLOLPAN 481
::: : |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Qy 358 YRNGDLLIDHSGAYVAQYYITWNELSYDHQGEVLTPRAMDRNQGDLTAHFTTSIPLKGN 417

Db 482 ARNIHVEGEGATGTAADPMW-TVINKKNLPVPHREIYLGTTINPWVEDNV 532
 ||:| | ||||: || | | : |||| | : |||| | |||:|
 QY 418 VNLNVKIRRECTGPAWE-WMRTVYEKTDLPVLRKRITSIWGTILYQVEEKV 468

RESULT 4
ID Q48772
C 048772
PRELIMINARY:
PRT; 50 AA.

DT	01-NOV-1996 (TREMBLREL, 01, CREATED)
DT	01-NOV-1996 (TREMBLREL, 01, LAST SEQUENCE UPDATE)
DT	01-NOV-1998 (TREMBLREL, 08, LAST ANNOTATION UPDATE)

0C BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
0S LISTERIA MONOCYTOGENES.
GN HLY.
02 BACILLOVIBRIOLIN U (FRAGMENT).
03

CC LISTERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-11984, TYPE 1;

RA RASMUSSEN O.F., SKOUBOE P., DONS L., ROSSEN L., OLSEN J.E.;

"Listeria monocytogenes exists in at least three evolutionary lines:
RT evidence from flagellin, invasive associated protein and

RT listeriolysin O genes."
RL MICROBIOLOGY 141:2053-2061(1995).

DR EMBL: X85855; G940601; -
FT NON_TER 1 50
FT NON_TER 1 50

SEQUENCE 50 AA: 5368 MW; 7080399E CRC32;

Query Match 3.8%; Score 124; DB 2; Length 50;
Best Local Similarity 34.7%; Pred. No. 1.83e-03;
Matches 17; Conservative 14; Mismatches 18; Indels 0; Gaps 0;

Db 2 NSSEKAVIYGGASANDQVOIIDNGLDRLDKKATFNRTPGVPIAYT 50
QY 284 NTEKAVILGGDPSSGARVYTKVMDVEDLIDGSRFTADHPGLISYT 332

RESULT 5 PRELIMINARY; PRT; 50 AA.

AC Q48773;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE LISTERIOLYSIN O (FRAGMENT).
GN HLY.
OS BACTERIA MONOCYTOGENES.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC LISTERIA.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-206.1.1, TYPE 2A;
RX MEDLINE: 96118685.

RA RASMUSSEN O.F., SKOUBOE P., DONS L., ROSSEN L., OLSEN J.E.;
RT "Listeria monocytogenes exists in at least three evolutionary lines:
RT evidence from flagellin, invasive associated protein and

RT listeriolysin O genes".
RL MICROBIOLOGY 141:2053-2061(1995).

DR EMBL: X85935; G940637; -
FT NON_TER 1 1
FT VARIANT 13 13 F -> Y.
FT NON_TER 50 50

SEQUENCE 50 AA: 5429 MW; A4B59721 CRC32;

Query Match 3.6%; Score 118; DB 2; Length 50;
Best Local Similarity 34.7%; Pred. No. 1.29e-02;
Matches 17; Conservative 14; Mismatches 18; Indels 0; Gaps 0;

Db 2 NSSEKAVIYGGASANDQVOIIDNGLDRLDKKATFNRTPGVPIAYT 50
QY 284 NTEKAVILGGDPSSGARVYTKVMDVEDLIDGSRFTADHPGLISYT 332

RESULT 6 PRELIMINARY; PRT; 450 AA.
AC Q68518;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 49.7 KD PROTEIN.
CS MYXOCOCCUS XANTHUS.
OS BACTERIA; PROTEOBACTERIA; DELTA SUBDIVISION; MYXOBACTERIA;
OC MYXOCOCCALES; CYSTOBACTERINEAE; MYXOCOCCACEAE; MYXOCOCCUS.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-DZFL;
RX MEDLINE: 96347136.

RA TRUDAU K.G., WARD M.J., ZUSMAN D.R.;
RT "Identification and characterization of FrzZ, a novel response
RT regulator necessary for swarming and fruiting-body formation in
RT Myxococcus xanthus".
RL MOL. MICROBIOL. 20:645-655(1996).

SEQUENCE 450 AA: 49686 MW; E3A6FC67 CRC32;

Query Match 3.5%; Score 113; DB 2; Length 450;
Best Local Similarity 26.8%; Pred. No. 6.29e-02;
Matches 30; Conservative 26; Mismatches 49; Indels 7; Gaps 6;

Db 9 PLVIVTSDMTFAPKL-LFADPKGRVMEHPYLLATLRSGEELVPPOD-KPIPLSTG-RLV 65
QY 233 PLVITL-SVYAVGQVYLKETTSSXSEVDAFAELIKYKAPQLEKQI-LDNTFVAV 290

66 HLPGRPLVGLHPETGELLEVREMKVGGKTFVNAVAGALLPGYTRTFLPGEV 117
QY 291 ILGDPSSGARVYTKVMDVEDLIDGSRFTADHPG--LPISYTSFLRDNV 340

RESULT 7 PRELIMINARY; PRT; 748 AA.

AC Q12098;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE CHROMOSOME XII READING FRAME ORF YLR135W.
GN L9606.6 OR L3140.
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCAROMYCETALES;
OC SACCAROMYCETACEAE; SACCAROMYCES.
RN [1]

RP SEQUENCE FROM N.A.
RC DELIUS H., HEBLING U.;
RT SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RX MEDLINE: 96118685.

RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,
RA FAVELL A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,
RA HALLSWORD K., HARKINS J., HILLIER L., JIER M., JOHNSON D.,
RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARIS E., MENEZES S.,
RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIKEN L., RILES L.,
RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOLDMAN P., VAUDIN M.,
RA WILSON R., WATERSTON R.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]

RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RX MEDLINE: 96118685.

RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,
RA FAVELL A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,
RA HALLSWORD K., HARKINS J., HILLIER L., JIER M., JOHNSON D.,
RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARIS E., MENEZES S.,
RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIKEN L., RILES L.,
RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOLDMAN P., VAUDIN M.,
RA WILSON R., WATERSTON R.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]

RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RX MEDLINE: 96118685.

RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,
RA FAVELL A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,
RA HALLSWORD K., HARKINS J., HILLIER L., JIER M., JOHNSON D.,
RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARIS E., MENEZES S.,
RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIKEN L., RILES L.,
RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOLDMAN P., VAUDIN M.,
RA WILSON R., WATERSTON R.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]

Query Match 3.5%; Score 112; DB 3; Length 748;
Best Local Similarity 18.8%; Pred. No. 8.58e-02;

QY	246	YYLLETTSSXSEVAF	EAL	FK	269
RESULT	13				
ID	087083	PRELIMINARY;	PRT;	1361	AA.
AC	087083;				
DT	01-NOV-1998	(TREMBLREL. 08, CREATED)			
DT	01-NOV-1998	(TREMBLREL. 08, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	S-LAYER PROTEIN.				
OS	CAMPYLOBACTER RECTUS.				
OC	BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;				
OC	CAMPYLOBACTER.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 33238;				
RX	MEDLINE; 98442421.				
RA	MIYAMOTO M., MAEDA H., KITANAKA M., KOKUGUCHI S., TAKASHIBA S.,				
RA	MURAYAMA Y.;				
RT	"The S-layer protein from Campylobacter rectus: sequence				
RL	determination and function of the recombinant protein.";				
RL	FEMS MICROBIOL. LETT. 166:275-281(1998).				
DR	EMBL; AB001876; D1034498; -.				
KW	S-LAYER.				
SO	SEQUENCE	1361	AA;	144904	MM; 55A9D736 CRC32;
Query Match		3.3%;	Score 106;	DB 2;	Length 1361;
Best Local Similarity		20.4%;	Pred. No. 5.33e-01;		
Matches	38;	Conservative	49;	Mismatches	89; Indels 10; Gaps 10;
Db	178	VDHMSNEHGGKAELETTNTDNTATNFANPMK-HNFGGIDRIMTLOSCKLIGDYSRHD	236		
QY	115	VEDPSNSVVRQAVNDLLAKWHODYGQVNNVPARQYEKITAHSME-OLKVKFGSDFEKGTG	173		
Db	237	NTLVEFGQANADEGDETSRTPLTNTQINIEYTGVTNLDL-RDSNDYEKINHRTK	295		
QY	174	NSLIDIDNSVHSGE-KOI-QIVNKKOI-YTTSVDAYKNGEDVFQDVTYEDLKQKQISA	230		
Db	296	EAGKKEFVESIG-OKLVGMRLANYAK-KDIDVKEH-KKGVLGSFEDSKSNVFLENVAKS	352		
QY	231	ERPLVY-ISKVAAYGRQYLYLKLSTXSXEVEAFEALIKGVKVAPOREWQIIDLNTLVKA	289		
Db	353	LSITSD	358		
QY	290	VILGSD	295		
RESULT	14				
ID	030524	PRELIMINARY;	PRT;	1361	AA.
AC	030524;				
DT	01-JAN-1998	(TREMBLREL. 05, CREATED)			
DT	01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	S-LAYER PROTEIN.				
GN	CRS.				
OS	CAMPYLOBACTER RECTUS.				
OC	BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;				
OC	CAMPYLOBACTER.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=314;				
RX	MEDLINE; 98187925.				
RA	WANG B., KRAIG E., KOLODRUBETZ D.;				
RA	"A new member of the S-layer protein family: characterization of the				
RT	crs gene from Campylobacter rectus.";				
RL	INFECT. IMMUN. 66:1521-1526(1998).				
DR	EMBL; AF010143; G245961; -.				
SO	SEQUENCE	1361	AA;	144385	MM; CAFE081F CRC32;
Query Match		3.3%;	Score 106;	DB 2;	Length 1361;
Best Local Similarity		20.4%;	Pred. No. 5.33e-01;		
Matches	38;	Conservative	49;	Mismatches	89; Indels 10; Gaps 10;

 W P E R E H
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Msearch_p protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Aug 30 16:14:04 1999; Maspar time 18.38 Seconds
 Tabular output not generated. 544.892 Million cell updates/sec

Title: >US-09-120-044-4
 Description: (1-471) from US09120044.pep
 Perfect Score: 3127
 Sequence: 1 MANKAVNDFILAMNYDXKTL.....TISWGTTLXPQYEDKVENND 471

Scoring table:
 PAM 150
 Gap 11

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseg35
 1:part1 2:part2 3:part3 4:part4 5:parts 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 35.164; Variance 168.905; scale 0.208

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	3127	100.0	471	1	R05923	Immunogenic pneumolys	1.49e-263
2	2961	94.7	471	1	R05924	Immunogenic pneumolys	1.08e-248
3	1387	44.4	480	6	R33842	Streptolysin O varian	1.86e-108
4	1369	43.8	480	6	R33841	Soluble streptolysin	7.25e-107
5	1362	43.6	571	2	R10376	Streptolysin O deriva	3.01e-106
6	1099	35.1	371	1	R06000	Bacteriophage lambda	4.69e-83
7	370	11.8	113	33	W60952	Streptococcus pneumon	4.16e-20
8	112	3.6	11	17	R76733	Listeriolysin O haemo	5.73e+00
9	108	3.5	1228	17	W23862	S-layer protein enco	1.06e+01
10	102	3.3	299	33	W60976	Bacillus stearothermo	1.06e+01
11	102	3.3	469	20	W04722	Streptococcus pneumon	2.63e+01
12	101	3.2	188	29	W55347	Aromatic acyl transfe	3.05e+01
13	101	3.2	418	29	W55528	H. pylori ORF 01ccl16	3.05e+01
14	100	3.2	860	8	R42962	H. pylori ORF 29ge303	3.54e+01
15	100	3.1	458	8	R41781	Bovine MTP	6.39e+01
16	96	3.1	458	8	R41781	Rabbit HBI protein wh	6.39e+01

17	97	3.1	606	18	R93121	Tilapia prolactin rec	5.52e+01
18	97	3.1	630	18	R93120	Tilapia prolactin rec <td>5.52e+01</td>	5.52e+01
19	98	3.1	667	22	W20753	H. pylori transporter <td>4.76e+01</td>	4.76e+01
20	98	3.1	1810	16	R94563	Chicken cytochrome <td>4.76e+01</td>	4.76e+01
21	94	3.0	114	2	R07691	Myobacterial 65KD ant <td>8.55e+01</td>	8.55e+01
22	94	3.0	355	12	R64226	Human 38kDa FK-506 bi <td>8.55e+01</td>	8.55e+01
23	94	3.0	527	33	W61278	Human monomamine oxida <td>8.55e+01</td>	8.55e+01
24	94	3.0	527	2	R05079	Human monomamine oxida <td>8.55e+01</td>	8.55e+01
25	93	3.0	625	23	W18008	Arbidopos CERI prot <td>8.55e+01</td>	8.55e+01
26	94	3.0	1057	36	W72067	HSV-2 strain SB5 Cont <td>8.55e+01</td>	8.55e+01
27	94	3.0	1114	36	W72066	HSV-2 strain SB5 Cont <td>8.55e+01</td>	8.55e+01
28	94	3.0	1114	36	W72065	HSV-2 strain SB5 Cont <td>8.55e+01</td>	8.55e+01
29	93	3.0	1786	23	W24790	P. falciparum liver s <td>9.88e+01</td>	9.88e+01
30	91	2.9	480	38	W74814	Human secreted protei <td>1.32e+02</td>	1.32e+02
31	91	2.9	480	38	W74814	Human secreted protei <td>1.32e+02</td>	1.32e+02
32	91	2.9	796	38	W85598	Cadherin-11. <td>1.32e+02</td>	1.32e+02
33	91	2.9	796	23	W25636	Human cadherin-11. <td>1.32e+02</td>	1.32e+02
34	91	2.9	796	23	W13134	Human MTP. <td>1.52e+02</td>	1.52e+02
35	90	2.9	894	8	R50007	Human MTP. <td>1.52e+02</td>	1.52e+02
36	90	2.9	964	3	R13618	C3 vegetable PEPC. <td>1.52e+02</td>	1.52e+02
37	92	2.9	3080	6	R35081	ZYMV polyprotein. <td>1.14e+02</td>	1.14e+02
38	88	2.8	84	32	W28284	Staphylococcus aureus <td>2.02e+02</td>	2.02e+02
39	88	2.8	176	22	W20335	H. pylori cytoplasmic <td>2.02e+02</td>	2.02e+02
40	89	2.8	211	1	R90412	Plasmodium falciparum <td>1.75e+02</td>	1.75e+02
41	88	2.8	531	22	W20739	H. pylori cytoplasmic <td>2.02e+02</td>	2.02e+02
42	89	2.8	649	38	W81977	Ehrlichia sp. B3 prot <td>1.75e+02</td>	1.75e+02
43	88	2.8	747	39	R49732	Sequence encoded by h <td>2.02e+02</td>	2.02e+02
44	89	2.8	747	39	W89585	Human ATP-binding cas <td>1.75e+02</td>	1.75e+02
45	89	2.8	757	20	W03179	Bovine poly-immunoglo <td>1.75e+02</td>	1.75e+02

ALIGNMENTS

RESULT 1
 ID R05923 standard; protein; 471 AA.
 AC R05923;
 DT 23-NOV-1990 (first entry)
 DE Immunogenic pneumolysin variant.
 KW Pneumolysin; vaccine; pneumonia; meningitis; bacteraemia; ds.
 OS Streptococcus pneumoniae.
 PN W09006951-A.
 PD 28-JUN-1990.
 PE 15-DEC-1989; AU0539.
 PR 15-DEC-1988; AU-001989.
 PA (PATO/) PATON J C.
 PI PATON JC, HANSMAN DJ, BOULINOIS GJ, ANDREW PM, MITCHELL TJ,
 PI WALKER JA;
 DR NPI: 90-224494/29.
 DR N-PSDB: 005270.
 PT New non-toxic, immunogenic mutants of pneumolysin - useful in
 PT protective vaccines against Streptococcus pneumoniae, and DNA
 PT sequences encoding them.
 PS Claim 5; Page 11; 25pp; English.
 CC Vaccines are non-toxic and antigenic to wild type pneumolysin,
 CC making them useful in vaccination against pneumonia and associated
 CC meningitis, bacteraemia etc.
 SQ Sequence 471 AA;
 Query Match 100.0%; Score 3127; DB 1; Length 471;
 Best Local Similarity 94.5%; Pred. No. 1.49e-263;
 Matches 445; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Db 1 markavndfilamnykklthggesianrfikenglpdfevlerkkrlstntdsi 60
 1 MANKAVNDFILAMNYDXKTLTHGGSINRFXKEGNOJPEFXVXKRRKSLSTNTSDI 60
 |||||||
 Db 61 svatndsrlypgallvveitllemptlajvdrapmtysilpplasadstlqyedsn 120
 1 svatndsrlypgallvveitllemptlajvdrapmtysilpplasadstlqyedsn 120
 |||||||
 Qy 61 xvatndsrlypgallvveitllemptlajvdrapmtysilpplasadstlqyedsn 120
 61 xvatndsrlypgallvveitllemptlajvdrapmtysilpplasadstlqyedsn 120
 |||||||
 Db 121 ssyrgavndllakwbqdygqvnvparmqyekltahsmegllkvkfgsfdektqnsldidf 180
 121 ssyrgavndllakwbqdygqvnvparmqyekltahsmegllkvkfgsfdektqnsldidf 180
 |||||||
 Qy 121 ssyrgavndllakwbqdygqvnvparmqyekltahsmegllkvkfgsfdektqnsldidf 180
 121 ssyrgavndllakwbqdygqvnvparmqyekltahsmegllkvkfgsfdektqnsldidf 180
 |||||||

D	181	nsvhsgkqldqivfkqjyrvsvdaunkngdyfcdtrvvedlkgtrgsaaerplvyissv	240
D	181	nsvhsgkqldqivfkqjyrvsvdaunkngdyfcdtrvvedlkgtrgsaaerplvyissv	240
O	181	nsvhsgkqldqivfkqjyrvsvdaunkngdyfcdtrvvedlkgtrgsaaerplvyissv	240
D	241	ayrgqvyllketskdsdeveaafaillkykvapqtcwqjldntevxavillggpssga	300
O	241	ayrgqvyllketskdsdeveaafaillkykvapqtcwqjldntevxavillggpssga	300
O	241	ayrgqvyllketskdsdeveaafaillkykvapqtcwqjldntevxavillggpssga	300
D	301	rvtvgtvnmvedlkegarfcbadpnpjisytsflndvatfgnsidvayektrayrn	360
O	301	rvtvgtvnmvedlkegarfcbadpnpjisytsflndvatfgnsidvayektrayrn	360
O	301	rvtvgtvnmvedlkegarfcbadpnpjisytsflndvatfgnsidvayektrayrn	360
D	361	gdlllhshgsaayaqyilwdelisydhgkvevlprkawdrnqgltahtftsipjkgvyn	420
O	361	gdlllhshgsaayaqyilwdelisydhgkvevlprkawdrnqgltahtftsipjkgvyn	420
O	361	gdlllhshgsaayaqyilwdelisydhgkvevlprkawdrnqgltahtftsipjkgvyn	420
D	421	lsvkifrectglawewrtvyekefdlprkrktisjwgttllpyrvckvnd	471
O	421	lsvkifrectglawewrtvyekefdlprkrktisjwgttllpyrvckvnd	471
O	421	lsvkifrectglawewrtvyekefdlprkrktisjwgttllpyrvckvnd	471

RESULT	2
ID	R05924 standard: protein; 471 AA.
AC	R05924;
DT	23-NOV-1990 (first entry)
DE	Immunogenic pneumolysin variant.
KW	Pneumolysin: vaccine; pneumonia; meningitis; bacteraemia; ds.
OS	Streptococcus pneumoniae.
FH	key Location/Qualifiers
FT	misc_difference 428
FT	/label=G, A, S
FT	misc_difference 435
FT	/label=Q, D
PN	W09006951-A.
PD	28-JUN-1990.
PF	15-DEC-1989; AU0539.
PR	16-DEC-1988; AU-001989.
PA	(PATO/) PATON J C.
PI	PATON JC, HANSMAN DJ, BOUNOIS GJ, ANDREW PM, MITCHELL TJ,
PI	WALKER JA;
DR	WPI; 90-224494/29.
PT	New non-toxic, immunogenic mutants of pneumolysin - useful in
PT	protective vaccines against Streptococcus pneumoniae, and DNA
PT	sequence encoding them.
PS	Disclosure; P: English.
CC	Vaccines are non-toxic and antigenic to wild type pneumolysin,
CC	making them useful in vaccination against pneumonia and associated
CC	meningitis, bacteraemia etc.
SQ	Sequence 471 AA;

```

Query Match 94.7% Score 2961 DB 1.96 Length 471:
Best Local Similarity 91.3% Pred. No. 1 08e-248:
NCBI Genbank 430: Conservative 5; Mismatches 36; Indels 0; Gaps 0;
Db 1 mankandfilamnydkkklthggsienfifikegnlpgdefvlerkxrsintnsdi 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 MANKAVNDPILAMNDYDKKLLTHGGSLENFIFIKENGLPGDEFVLERKXRSINTNSDI 60
Luo 61 svlatndsrlypgallvwdecllemptllavdrpmtytsidpglaasdsfiyepdsn 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 XYXATXDSRLYGCALLVDETELKNNPPLADRAPMYXXLLPGLASDSFIVDEPDSN 120
Qy 61 XYXATXDSRLYGCALLVDETELKNNPPLADRAPMYXXLLPGLASDSFIVDEPDSN 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 ssvrgaxvdlakwphdvgvgvnnvparmgvkekthsmegklvkfgsdfeftgnsldidf 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 121 SSVRGAXXDLKMHODVGVQVNNVPRKXQYKEXKTHSHSDEQLVKFGSGFEKXGNSLDIDF 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 nsyhsgskgldqivnfkqilytyvsavavkpgvgfdgtlvvedlkgqrgisaerplvyvysv 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 181 NSYHSGSKGIQIVNFKQIYLYTVSVAVVKPGVGFDTLVVEDLKGRIASERPLVYISXV 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 aysgrvyllkletsdsdevaafealixkvxvapgtewkqllidntevkavvylggdpsga 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

```
QY 241 AXRRVYLKLETTSSXSEVBAFEALIKGVKVAPOJEMQJLJNDTXKAVLJGSDPSSGA 3000
Db 301 rrvvgkvdmvedlllqegsrffadhpqlpisytsfirhvatfmscdyektvetyrn 3600
QY 301 RVTGSKVMVEDLLIOESRFTADHPGRLPSYTSFELRDWVAPEONSIDYETIETVATARN 3600
Db 361 gdlldrgavvaqyyltfldelsfnbgkxevltpkxfidngsdlahtfsjplkqavrn 4200
QY 361 GDLLDHSGAVVAOYIYTWKXELSDHOGKXVLLPDKWMDNGDLSAHTTSLPLKGNVRN 4200
Db 421 lsvkkregjlaqfgrfvtvyeoktdlpvkrkrtssqgtllpyqvdekvend 471
QY 421 LSVKIRECTGLAMEMVRVYEKNDLXVRRRTISMGTLTPQVEDKVEND 471
```

ID	RESULT	3
AC	R3842	standard; Protein; 480 AA.
DT	R3842	(first entry)
DE	Streptolysin O variant mSLO.3/6.	
KM	SLO; soluble; haemolytic activity; wild type; anti-SLO antibodies;	
KM	Streptococcus pyogenes; ASO.	
OS	Streptococcus pyogenes.	
PN	W09305155-4.	
PD	18-MAR-1993.	
PE	03-AUG-1992; U06380.	
PR	30-AUG-1991; US-752428.	
PA	(BECI) BECKMAN INSTR INC.	
PI	Adams CW;	
DR	WPI: 93-100979/12.	
DR	NPSDB; Q38287.	
PT	Streptolysin O variants produced by recombinant DNA technology -	
PT	having no haemolytic activity and recognised by wild-type	
PT	anti-streptolysin O antibodies, useful in diagnosis of	
PT	Streptococcus pyogenes infection	
PS	Claim 17; Fig 2: 80pp; English.	
CC	The sequence represents a soluble variant of Streptolysin O (SLO)	
CC	having no haemolytic activity. The variant SLO can specifically	
CC	recognise wild-type anti-SLO antibodies (ASO) but has no haemolytic	
CC	activity. It can be used for the determination of previous and/or	
CC	current infection by Streptococcus pyogenes. It enables ASO assays	
CC	to become commercially viable and avoids the risk of handling haemo-	
CC	lytically active S. pyogenes wild-type SLO.	
CC	Sequence	480 AA;
Q9		

[illegible]

/label=Iamdba N promoter - SLO.

```
FT CA2003307-A.
PD 18-MAY-1990.
PR 17-NOV-1989; 026008.
PR 18-NOV-1989; GB-027038.
PA 17-NOV-1989; GB-026008.
PA (KEHO//) KEHOM M.
PI KEHOM M:
PI WPI: 90-217310/29.
DR N-FSDS: Q05240.
PT Non-toxic and non-cytolytic derivs. of streptolysin O - used for
PT raising antibodies, purifying antibodies or detecting antibodies
PS to streptolysin O.
PS Disclosure: PP; English.
CC SLO derivative is epitopic to its wild type parent, and may be
CC used in detecting anti-SLO Abs, indicating presence of Streptococcus
CC pyogenes.
SO Sequence. 371 AA;

Query Match          35.1%; Score 1099; DB 1; Length 371;
Best Local Similarity 42.6%; Pred. NO. 4.69e-83;
Matches 147; Conservative 74; Mismatches 124; Indels 0; Gaps 0;

Db      20 aanpllyvsakpvngwhnyspgntlparqtgytesmynksqsgleaainvnsklldgtllg 79
       ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
vY      118 psnsvrARAXXXDLKMLKHQDVGQVVNNVAPRXYEKETXAHNSMEOLKVKGSGDFEKKXSND 177
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Z      60 idfksisxgekkvmaiaeykqifvytsanljpmnadvfdksvtfkelqrkyvsneapplfv 139
       |||::: |::: |::: |::: |::: |::: |::: |::: |::: |::: |::: |::: |:::
Z      173 IDFNsvHSEKKIQIVNKKQIIITYTSVDVAKNPGDVFGDVTYVEDLKRGRISAEPLVIYI 237
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      140 snvaygrtvfkletssksndveaafsaalqgdvtktngkysdlensftavylggdaa 199
       | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      238 SXVAAXXROYLLELTETSSXSEVEAFELIKGVKYPOTEKQOIIDNTYKAVILIGDPS 297
       | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
UW      200 ehmkvtdfdyirnvikdnatfsrknpapisytsvflknkjiagvnrtreyeltste 259.
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      298 SGARVTVGVDWEEDLIQGSNFRTADHDELPYSTTSLRLDNVAAITPONSIDYEATKYTA 357
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      260 ytsqgnishbgayvaagyeilwdeainyddkgvkvtkrrwdnmwyskspsfsvlajgan 319
       | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      358 YRRGDLLDHSGAAYAQQYITWXELSYDHOGKEVTELPKAMPBNGODTLFAHFTTSIPLKGN 417
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      320 sgnrirmarectglawewwrkyiderdvylskelnvnmsgstlsp 364
       ||::: :|||:|||:|||::: : : : : : : : : : : : : : : : : : : :
QY      418 VRNLsvKIRECTGLAWEMRWRYERTDLVLVKRKRTISlgTtLYP 462

RESULT 7
ID M60952 standard; Protein: I13 AA.
AC W60952;
DT 13-OCT-1998 (first entry)
DE Streptococcus pneumoniae encoded polypeptide.
KM coding region; ORF; Open reading frame; antibacterial;
KW Infection; Prevention; meningitis.
OS Streptococcus pneumoniae.
PN MO9819689-A1.
PD 14-MAY-1998.
PD 27-OCT-1997; 019226.
PR 01-NOV-1996; US-029930.
PA (SMIR ) SMITHKLIN BECHAM CORP.
PA (SMIR ) SMITHKLIN BECHAM PLC.
PI Black Mt, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO,
PI Reid RH, Zarfos PW;
PI WPI: 98-286586/25.
DR N-PDSB: V37352.
PT New isolated nucleic acids from Streptococcus pneumoniae - useful,
PT e.g. for identifying anti-bacterial(s) for treatment and prevention
PT of meningitis
PS Claim 11: page 53: 130pp; English.
CC The sequence is that of the polypeptide encoded by a region isolated
CC from S. pneumonia which shows homology to pneumolysin.
CC The protein, or agonists of it, may be useful as an antibacterial for
```

	CC	treatment or prevention of infection, specifically caused by S.
	CC	pneumoniae (particularly meningitis) but possibly also Helicobacter
	CC	pylori (ulcers and gastric cancer). It may be of particular
	CC	use before insertion of an in-dwelling device or any other
	CC	invasive procedure. The protein, or nucleic acid encoding
	CC	it, can also be used in vaccines to induce a cellular
	CC	and/or humoral immune response, or to screen for other
	CC	antibacterials. The DNA may also contain flanking sequences
	CC	that are potential sources of control elements for bacterial
	CC	gene expression. Detecting a sequence encoding the protein
	CC	can be used diagnostically, e.g. to detect a mutation for
	CC	serotyping or classifying infectious agents.
SQ	Sequence	113 AA;
Db	Query Match	11.8%; Score 370; DB 33; Length 113;
OY	Best Local Similarity	43.2%; Pred. No. 4,16e-20;
	Matches	48; Conservative 31; Mismatches 31; Indels 1; Gaps 1;
Db	2 kdkllilqhsayaiarsltweevyvdkgdvvrshswegnngtavgvlnlpikenn	61
OY	359 RRGDLDDSGAVAAQXYITWTXELSYDHQGKEVLPFRKARDNRGODTLAFTTSPILKNMV	418
	:: :::: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :	
Dd	62 nrlvyklettqtgllwnrtgtlyenrrplaqphkkihwftlnskrsddv	112
OY	419 RLRLSKIRECTGLAMEMWRIVTEKTDLTVRKRTISIMWGITLIPDY-EDKV	468
	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :	
	RESULT 8	
ID	R88135 standard; peptide: 11 AA.	
AC	R88135.	
DT	29-AUG-1996 (first entry)	
DE	Listeriolysin O haemolytic domain; streptolysin O; pneumolysin;	
KM	Listeriolysin O haemolytic; domain; streptolysin O; pneumolysin;	
KW	soluble molecular complex; specific; targeted; polynucleotide;	
KW	endosome lysate; bacterial component; cytoplasm; therapeutic;	
OS	In vivo; ex vivo; in vitro; delivery; cell specific; hemolytic.	
PN	MO9600792-A1.	
PD	11-JAN-1996.	
PF	27-JUN-1995; UO8091.	
PR	29-JUN-1994; US-267710.	
PR	07-JUN-1995; US-484009.	
PA	(TARGET-) TARGETECH INC.	
PT	(UYCO-) UNIT CONNECTICUT.	
PI	Carmichael E, Spitalny GL, Wu CH, Wu GY, Zhang Y;	
DR	WPL: 96-077502/08.	
PT	Soluble molecular complex for specific targeting of PN to cell -	
PT	internalises polynucleotide (PN) into endosome, then lyses endosome	
PS	to release PN into cell's cytoplasm	
PS	Example 3; Page 13; 23pp; English.	
CC	The present peptide is a listeriolsyin O (LLO) haemolytic domain	
CC	(HD) fragment which is 100% conserved in the LLO homologues	
CC	streptolysin O, and pneumolysin A. A 201 bp fragment encoding the LLO	
CC	HD can be used instead of the full length LLO coding sequence in a	
CC	soluble molecular complex, for the specific targeting of a	
CC	polynucleotide (PN) to a cell. The complex comprises a PN, a	
CC	carrier comprising a PN binding agent and a cell specific binding	
CC	agent, which binds to a surface mol. of the cell and its	
CC	internalised into an endosome, and a bacterial component (i.e. the	
CC	201 bp LLO HD fragment), which lyses the endosome to release the	
CC	PN into cell's cytoplasm. The complex can be used therapeutically	
CC	for enhanced in vivo, ex vivo or in vitro delivery of a PN to a	
CC	specific cell.	
SQ	Sequence	11 AA;
Dd	Query Match	3.6%; Score 112; DB 17; Length 11;
OY	Best Local Similarity	100.0%; Pred. No. 5.73e+00;
	Matches	11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	427 ECTGLAMEWMR	437
	1 ectglamewmr	11


```

RESULT 9
ID R7673 standard; Protein; 1228 AA.
AC R7673;
DE 16-JUL-1996 (first entry)
KW S-layer protein encoded by sbsA gene.
OS Bacillus stearothermophilus.
FH Key
FT peptide
FT protein
FT /label= sig_peptide
FT /label= mat_protein
DE425527-A1.
PD 25-JAN-1996.
PR 19-JUL-1994; 425527.
PR 19-JUL-1994; DE-425527.
PA (VOGE-) VOGELBUSCH GMBH.
PI Lubitz W;
DR WPI: 96-077933/09.
DR N-PSDB: T08695.
PT Nucleic acid encoding signal peptide of Bacillus stearothermophilus
PT S-layer protein - which has a lysine content of at least 10 per
PS cent
PS Disclosure: Page 6-7; 12pp; German.
CC The claimed signal peptide encoding sequence (see tag e of T08695)
CC is pref. operably linked at the 3' end to a protein encoding
CC sequence. The protein is pref. the S-layer protein (see tag f of
CC T08695). At the 5' end, the signal peptide encoding sequence is
CC pref. linked to an expression control sequence, pref. the sequence
CC given in T08696.
SQ Sequence 1228 AA;

Query Match 3.5%; Score 108; DB 17; Length 1228;
Best Local Similarity 23.5%; Pred. No. 1.06e+01;
Matches 39; Conservative 43; Mismatches 74; Indels 10; Gaps 10;

DB 200 ydliv-amkar-ev-gdavkagnldkakaavdginqylpkvtdakfktelt-evakkalda 255
QY 200 YTVSVDVAVKNGDVFQDTVTVEDLKRGISAERPLVYSXVAVXKQVYLKLETTXSXEV 259
DB 255 deaaltpkvesvs-aintqkaveltavpvgnt-lklqsaanedvntvtriykvdg 313
QY 260 -EAAFEALIRGVKAVPQTEMKQI-LDNTXKAVILGGDPSSGARVYTGKDMVEDLIQEG 317
DB 314 nlpfalntadvslstdgkltvdastpfenteykvvvgkikdng 359
QY 318 S-RETDHPGLPISTYTSFLRDNVAVTFQNSTDY-VETKVTAYRNG 361

RESULT 10
ID W22862 standard; Protein; 1228 AA.
AC W22862;
DE 27-FEB-1998 (first entry)
DE Bacillus stearothermophilus S-layer protein sbs-A.
KW S-layer: sbs-A; vaccine; adjuvant; carrier; hybridisation assay;
KW molecular spinning nozzle; molecular laser.
OS Bacillus stearothermophilus.
FH Key
FT peptide
FT protein
FT /label= sig_peptide
FT /label= mat_peptide
DE19603649-A1.
PD 07-AUG-1997.
PR 01-FEB-1996; 003649.
PR 01-FEB-1996; DE-003649.
PA (LUBITZ/) LUBITZ W.
PA (SLEY/) SLEYTR U.
PI Kuen B; Lubitz W; Sleytr U;
DR WPI: 97-394558/37.
DR N-PSDB: T75487.
PT Preparation of S-layer proteins by expressing sbs-A gene in Gram

```

```

PT negative bacterium - or new sbs-B gene in any host, also new
PT recombinant proteins containing heterologous inserts, e.g.
PT epitope(s), useful as vaccines and adjuvants
PS Claim 1: Pages 9-14; 31pp; German
CC The present sequence is the Bacillus stearothermophilus PV72
CC S-layer protein, sbs-A. S-layer structures can be used as vaccines
CC or adjuvants, particularly when they include a bacterial ghost that
CC may contain additional epitopes in its membrane. Other uses of
CC recombinant sbs-A, depending on the nature of the inserted peptide,
CC are as an universal carrier for biotinylated reactants for use in
CC immunological or hybridisation assays (the insert is streptavidin),
CC to induce immune responses (epitopes), as a reagent for removing
CC cytokine or toxin from serum (antigenic epitopes), as a molecular
CC spinning nozzle (polyhydroxybutyrate synthase) and as a molecular
CC laser (luciferase).
SQ Sequence 1228 AA;

Query Match 3.5%; Score 108; DB 25; Length 1228;
Best Local Similarity 23.5%; Pred. No. 1.06e+01;
Matches 39; Conservative 43; Mismatches 74; Indels 10; Gaps 10;

DB 200 ydliv-amkar-ev-gdavkagnldkakaavdginqylpkvtdakfktelt-evakkalda 255
QY 200 YTVSVDVAVKNGDVFQDTVTVEDLKRGISAERPLVYSXVAVXKQVYLKLETTXSXEV 259
DB 255 deaaltpkvesvs-aintqkaveltavpvgnt-lklqsaanedvntvtriykvdg 313
QY 260 -EAAFEALIRGVKAVPQTEMKQI-LDNTXKAVILGGDPSSGARVYTGKDMVEDLIQEG 317
DB 314 nlpfalntadvslstdgkltvdastpfenteykvvvgkikdng 359
QY 318 S-RETDHPGLPISTYTSFLRDNVAVTFQNSTDY-VETKVTAYRNG 361

RESULT 11
ID W60976 standard; Protein; 299 AA.
AC W60976;
DE 13-OCT-1998 (first entry)
DE Streptococcus pneumoniae encoded polypeptide.
KW coding region; ORF; open reading frame; antibacterial;
KW infection; prevention; meningitis.
OS Streptococcus pneumoniae.
PN W09819689-A1.
PD 14-MAY-1998.
PR 27-OCT-1997; U19226.
PR 01-NOV-1996; US-029930.
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PIC.
PI Black MT, Hodgson JE, Knowles DTC, Lonetto MA, Nicholas RO,
PI Reid RH, Zarfos PN;
DR WPI: 98-286586/25.
DR N-PSDB: V37371.
PT New isolated nucleic acids from Streptococcus pneumoniae - useful,
PT e.g. for identifying anti-bacterial(s) for treatment and prevention
PT of meningitis
PS Claim 11, page 71-72; 130pp; English.
CC The sequence is that of the polypeptide encoded by a region isolated
CC from S. pneumoniae which shows homology to nucleotide-nucleotide pyro-
CC phosphorylase. It, or agonists of it, may be useful as an anti-
CC bacterial for treatment or prevention of infection, specifically caused
CC by S. pneumoniae (particularly meningitis) but possibly also Helicobacter
CC pylori (ulcers and gastric cancer). It may be of particular
CC use before insertion of an in-dwelling device or any other
CC invasive procedure. The protein, or nucleic acid encoding
CC it, can also be used in vaccines to induce a cellular
CC and/or humoral immune response, or to screen for other
CC antibacterials. The DNA may also contain flanking sequences
CC that are potential sources of control elements for bacterial
CC gene expression. Detecting a sequence encoding the protein
CC can be used diagnostically, e.g. to detect a mutation for
CC serotyping or classifying infectious agents.
SQ Sequence 299 AA;

```

	Query Match:	3.3%;	Score 102;	DB 33;	Length 299;	
	Best Local Similarity	20.4%;	Pred. No. 2.63e+01;			
	Matches	21;	Conservative	33;	Mismatches 43;	Indels 6; Gaps 6;
Dc	38 dycnaafdhbgqgkxvlfakeagvlag-ltvfgyftlfdaevtfqnphqtkdgrls	96	:	:	:	:
Cy	14 NYDXKKLLTHQGEIERFKKEGNQLPEFVXXEKRRKSLSITMS-DIXXATXSRLYP	72	:	:	:	:
Dc	97 gdl-vl-el-igsvrslldtoervallnqlghsglaamtayev	136	:	:	:	:
Cy	73 GALLVDETKLENNPTLLADVRAPMYTSXXXPLGLAS-SDSFLQ	114	:	:	:	:
RESULT	12					
ID	W04722; standard; Protein; 469 AA.					
AC	W04722;					
DT	06-FEB-1997 (first entry)					
DE	Aromatic acyl transferase of Gentiana triflora var. japonica.					
KW	Aromatic acyl transferase; transformation; anthocyanin pigment;					
KW	plants; acylation; colour; tone; colouration; colour change;					
KW	Gentiana triflora; Perlmia hybrida; Pellaea colimoides;					
KW	Senecio cruentus; Lavandula angustifolia.					
OS	Gentiana triflora var. japonica (Clone pGRF4).					
PN	W09625500-AI.					
PF	22-AUG-1996.					
PR	16-FEB-1996; J00348.					
PR	17-FEB-1995; JP-067159.					
PR	29-JUN-1995; JP-196915.					
PR	30-JAN-1996; JP-046534.					
PA	(SUNKR) SUNTORY LTD.					
PI	Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;					
PI	Nakao M, Tanaka Y, Yonekura K;					
DR	WP1; 96-393401/39.					
DI	N-PSSB; T37308.					
PT	DNA coding for aromatic acyl transferase - for transforming plants					
PT	e.g. of flowers					
PS	Claim 4; Page 53-57; 94pp; Japanese.					
CC	Vectors containing DNA fragments encoding proteins of plant origin					
CC	with aromatic acyl transferase activity may be used to transform					
CC	plants which produce anthocyanin pigments. The aromatic acyl					
CC	transferase acylates the pigments in the flower resulting in colour					
CC	tone changes and allowing new colourations to be produced. Six					
CC	specific DNA sequences encoding aromatic acyl transferase from					
SQ	different plants are described in T37308-T37313.					
	Sequence 469 AA;					
	Query Match	3.3%;	Score 102;	DB 20;	Length 469;	
	Best Local Similarity	19.8%;	Pred. No. 2.63e+01;			
	Matches	32;	Conservative	49;	Mismatches 70;	Indels 11; Gaps 11;
Dc	94 pkfygsrdgdstltivaesdqdflyghnql-vsndhglffymprvitmtqdy-kv-	150	:	:	:	:
Oy	146 ARQYQKRYKAHSHEDQKRVFGSDFE-KXGNSLDIDFNVSHESEKXI-QIVNKKQIYYTVS	203	:	:	:	:
Dc	151 iplvavqvrvfpmrjgavlatlhshadsksfvmfinawayank-fgk-dedllsanlip	208	:	:	:	:
Oy	204 VDAVKRPBGVPFDQTYVEDLK-QRGISARPLV-YISXAAYKROYLYKLLETSTXSXEVEA	261	:	:	:	:
Dc	209 sfdrslklvgljeetfwne-mgdvlemfsrfgkypfrinkv	249	:	:	:	:
Oy	262 AFE-ALIKGVKAPQTEMRKOILDNTFYKAVILLIGDPSGSGARV	302	:	:	:	:
RESULT	13					
ID	W55347; standard; Protein; 188 AA.					
AC	W55347;					
DT	17-JUN-1998 (first entry)					
DE	H. pylori ORF O1ccl6161orf11 protein.					
KW	Cytoplasmic; vaccine; prevention; treatment; infection; envelope;					
KW	identification; binding compound; bacteria; life cycle; activator;					
KW	Inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.					
OS	Helicobacter pylori.					

MO91737044-A1.
09-OCT-1997.
27-MAR-1997; U05223.
06-DEC-1996; US-761318.
29-MAR-1996; US-625811.
02-APR-1996; US-758731.
25-OCT-1996; US-736905.
28-OCT-1996; US-738859.
(ASTR) ASTRA AB.
Alm RA, Smith D;
WPI: 97-503122/46.
N-PSDB: V24756.
Helicobacter pylori nucleic acid sequences and encoded
polypeptide(s) - useful in vaccines to treat or prevent H. pylori
infection and for diagnosis of H. pylori infection
PS Claim 14, Page 571; 1145pp; English.
This sequence is a H. pylori protein of unspecified function.
CC The protein may be used in a vaccine to prevent or treat H. pylori
infection or to identify H. pylori life cycle activators or inhibitors. The
useful as potential H. pylori life cycle activators or inhibitors. The
DNA and probes derived from it may be used for the identification of
H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
acid sequences complementary to the DNA act as antisense sequences and
can be used to prevent the translation of H. pylori mRNA. Antibodies
against the protein can be used in immunoassays to evaluate the abundance
and distribution of H. pylori-specific antigens. The genomic sequence of
H. pylori (ATCC 55679) was determined from overlapping contigs generated
by mechanically shearing the bacterial DNA. The sequences were analysed
for ORF of at least 180 nucleotides and the predicted coding regions
defined by computer evaluation. To identify likely H. pylori antigens for
vaccine development, the amino acid sequences predicted from various ORF
were analysed for significant homology to other known or exported
membrane proteins. Having identified and determined the sequences of
interest, particular regions can be isolated from H. pylori by PCR
amplification for recombinant polypeptide production, e.g. in E. coli
hosts.
Sequence 188 AA;

	Query Match	3.2%	Score 101;	DB 29;	Length 188;
	Best Local Similarity	32.8%;	Pred.	No.3.0E+01;	
Matches	20;	Conservative	20;	Mismatches 18;	Indels 3; Gaps 3;
Dd	95 psslnrdsvsledskkr-lnaikdlifnknkaf-rgqlklneplkal-veaqkdgefk 151 : - - - : : : : : : : : : : : : : : : 210 PEDVQDITYEDLKGQRISAEPFLVISYAKRYLIKETTSXSEVEAFEAALIKG 269				
Dd	152 1 152				
Gy	270 V 270				
RESULT 14					
ID W55528 standard; Protein; 418 AA.					
AC W55528;					
DT 02-JUL-1998 (first entry)					
DE H. pylori ORF_29g3032L24336712.fl.5 cellular protein.					
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope; KW identification; binding compound; bacteria; life cycle; activator; KM inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis. OS Helicobacter pylori. PN MO9737044-A1. PD 09-OCT-1997 PF 27-MAR-1997 ; U053223. PR 06-DEC-1996; US-761318. PR 29-MAR-1996; US-625811. PR 02-APR-1996; US-758731. PR 25-OCT-1996; US-736905. PR 28-OCT-1996; US-738859. PA (ASTR) ASTRA AB. PI Alm RA, Smith D; DR MPI; 97-503122/46. DR N-PDOB: V24937. PT Helicobacter pylori nucleic acid sequences and encoded					

PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claims 14,93; Page 730-731; 1145pp; English.
 CC This sequence is a H. pylori cellular protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SZ Sequence 418 AA;

Query Match 3.2%; Score 101; DB 29; Length 418;
 Best Local Similarity 32.8%; Pred. No. 3.05e+01;
 Matches 20; Conservative 20; Mismatches 18; Indels 3; Gaps 3;

DB 382 1 382
 QY 210 PGDVRDITVTEDEKRGISAEPLVYISVAYXQVYIKLETTSSXKEVEAFALIKG 269
 VY 270 V 270

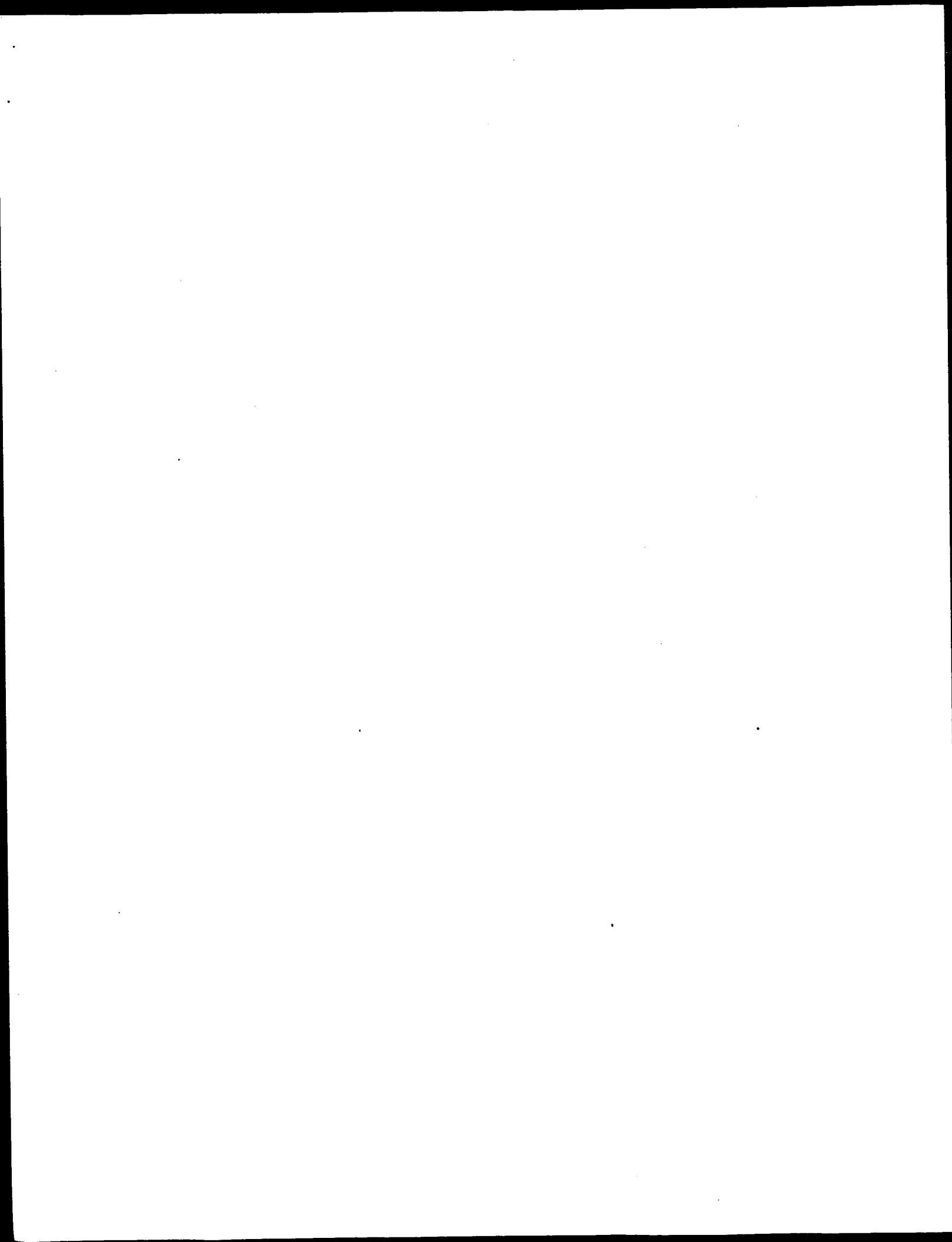
RESULT 15
 ID R42962 standard; Protein; 860 AA.
 AC R42962;
 DT 10-MAY-1994 (first entry)
 DE Bovine MTP.
 KW MTP; microsomal triglyceride transfer protein; obesity;
 KW atherosclerosis; detection; inhibitor; prevention; stabilisation;
 KW regression; lipid; serum cholesterol; SC; triglyceride; TG;
 KW phosphatidyl choline; PC; cholesterol ester; CE;
 KW pancreatitis; hyperglycemia.
 OS Bos taurus.
 PN AU0334064-A.
 PD 09-SEP-1993.
 PF 09-MAR-1993; 034064.
 PR 06-MAR-1992; US-847503.
 PA (SQUI) SQUIB & SONS INC E R.
 PI (WETT/) WETTERAU J R.
 PI Gregg RE, Sharp DY, Wetterau JR;
 DR WPI: 93-337295/43.
 DR N-PSDB: 050429.
 PT Nucleic acid sequence encoding sub-unit of MTP - used to express
 PT polypeptide for therapeutic use e.g. in reduction of obesity
 PT atherosclerosis etc.
 PS Claim 24; Page 25; 112pp; English.
 CC The sequences given in 050429-33 may be used for the detection
 CC of MTP or MTP inhibitors.
 CC Bovine and human MTP are useful for the prevention, stabilisation or
 CC regression of atherosclerosis or for lowering the level of serum
 CC lipids such as serum cholesterol, triglyceride, phosphatidyl choline
 CC or cholesterol ester in a mammalian species.
 CC Diseases associated or affected by serum lipid levels, such as
 CC pancreatitis, hyperglycemia, obesity, etc., may be treated.
 SQ Sequence 860 AA;

Query Match

3.2%; Score 100; DB 8; Length 860;

Best Local Similarity 20.1%; Pred. No. 3.54e+01;
 Matches 32; Conservative 38; Mismatches 81; Indels 8; Gaps 6;
 DB 481 ftdvkkttmrrlyhgnrklnkvtvrtaaillknpsymeknlls:igelpkemy 540
 QY 43 FVXKRRKRSLSTNTSDIXVAXIDSRLYPALLVDETXLENNPTLLAVRAPMYSSX 102
 DB 541 m--lsivgdllrfetpaskmrvylykemvahnvdrfsksgssaytgyvrtstasys 598
 QY 103 LPGLASSDSFLQVEDPSSSVRGAXDPLAKWHDYGVQVNNVFRKXOYEKXIAHSMQ-- 160
 DB 599 ldllysgsgllrrsn-lni-fgyiektcplngivieaq 635
 QY 161 LKVF-GSDFKXGNSDIDFNSVHSGE-KXIOIVNKKQ 197

Search completed: Mon Aug 30 16:15:24 1999
 Job time : 80 secs.



 WISE (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mparch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Aug 30 16:12:54 1999; Maspar time 20.74 Seconds
 Tabular output not generated. 909,898 Million cell updates/sec

Title: >US-09-120-044-4
 Description: (1-471) from US09120044.pep
 Perfect Score: 3127
 Sequence: 1 MANKAVNDFILAMNDXXKL.....TISMGITLYPQVEDKEND 471

Scoring table: PAM 150
 Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: p1r60
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 50.051; Variance 108.607; scale 0.461

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	3127	100.0	471	2	A28568	0.00e+00
2	1556	49.8	500	2	B43577	3.55e-265
3	1475	47.2	529	2	A43505	1.78e-249
4	1468	46.9	529	2	S24231	4.04e-248
5	1460	46.7	528	2	S22341	1.43e-246
6	1416	45.3	501	2	A37858	4.66e-238
7	1414	45.2	530	2	S22340	1.14e-237
8	1379	44.1	571	2	A43507	6.61e-231
9	1298	41.5	485	2	I39863	2.84e-215
10	464	14.8	112	2	S47298	3.51e-58
11	209	6.7	96	2	S47297	1.01e-14
12	113	3.6	265	2	S60947	2.77e-01
13	112	3.6	527	2	S03974	3.63e-01
14	112	3.6	744	2	A70385	3.63e-01
15	114	3.6	1361	2	T03415	2.11e-01
16	109	3.5	218	2	F64461	8.09e-01
17	108	3.5	462	1	S00552	1.05e+00
18	108	3.5	1228	2	I40468	1.05e+00
19	107	3.4	364	2	JC1188	1.37e+00
20	106	3.4	426	2	A70444	1.78e+00
21	105	3.4	776	1	RGECA8	2.30e+00
22	105	3.4	780	1	S77652	2.30e+00
23	102	3.3	365	2	F71665	4.95e+00

24	103	3.3	446	2	E69899	L-amino acid oxidase	3.84e+00
25	102	3.3	514	1	DNEB7T	threonine dehydratase	4.95e+00
26	103	3.3	722	2	S64492	hypothetical protein	3.84e+00
27	102	3.3	878	2	F64425	valine--tRNA ligase (4.95e+00
28	99	3.2	286	2	S55901	probable tagatose 1,6	8.16e+00
29	100	3.2	287	2	C71207	hypothetical protein	8.16e+00
30	100	3.2	349	2	H64456	hypothetical protein	8.16e+00
31	100	3.2	362	2	B70876	probable aminotransfe	8.16e+00
32	99	3.2	366	2	E69375	ABC transporter, ATP-	1.04e+01
33	101	3.2	418	2	G71952	hypothetical protein	6.36e+00
34	100	3.2	427	2	S75210	glycine hydroxymethyl	8.16e+00
35	101	3.2	501	2	A69415	conserved hypothetical	6.36e+00
36	100	3.2	502	2	UQ2341	arginine decarboxylas	8.16e+00
37	100	3.2	511	2	S58322	nucleolar protein NOP	8.16e+00
38	99	3.2	877	2	S58824	probable membrane pro	1.04e+01
39	100	3.2	887	2	A46764	microsomal triglyceri	8.16e+00
40	100	3.2	899	2	A44879	retinoblastoma protei	8.16e+00
41	101	3.2	1026	2	A49750	beta-lactamase (EC 3.	6.36e+00
42	99	3.2	1035	2	A64686	cation efflux system	1.04e+01
43	100	3.2	1123	1	WMDBH7	UL37 protein - human	8.16e+00
44	99	3.2	3027	2	JQ1917	polyprotein - parvini	1.04e+01
45	100	3.2	5762	2	A41819	proline-rich peptides	8.16e+00

ALIGNMENTS

RESULT 1

ENTRY 1
 TITLE A28568 #type complete
 ORGANISM pneumolysin - Streptococcus pneumoniae
 DATE 19-Nov-1988 #sequence_rev19-Nov-1988 #text_change

ACCESSIONS
 REFERENCE A28568
 #authors Walker, J.A.; Allen, R.L.; Falmagne, P.; Johnson, M.K.;

#journal Infect. Immun. (1987) 55:1184-1189
 #title Molecular cloning, characterization, and complete nucleotide
 sequence of the gene for pneumolysin, the
 sulphydryl-activated toxin of Streptococcus pneumoniae.

#cross-references M01D:87193109
 #accession A28568
 #molecule_type DNA

REFERENCE
 #residues 1-471 #label WAL
 #authors Mitchell, T.J.; Mendez, F.; Paton, J.C.; Andrew, P.W.;

#journal #journal
 #title Comparison of pneumolysin genes and proteins from
 Streptococcus pneumoniae types 1 and 2.
 #cross-references M01D:90326546
 #accession S12829
 #molecule_type DNA

CLASSIFICATION
 #residues 1-471 #label MT
 #cross-references EMBL:X52474; NID:947403; PID:947404
 #superfamily dipeptide transport protein

SUMMARY
 #product pneumolysin #status predicted #label MAT
 #length 471 #molecular_weight 52899 #checksum 8274

Query Match 100.0%; Score 3127; DB 2; Length 471;
 Best local similarity 94.5%; Pred. No. 0.00e+00;
 Matches 445; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

DB	1	MANKAVNDFILAMNDXXKLTHOGESIEIRKEGNOLEPEFVIRKRSTSTNSDI	60
QY	1	MANKAVNDFILAMNDXXKLTHOGESIEIRKEGNOLEPEFVIRKRSTSTNSDI	60
DB	61	SYTANDSRLPGALVYDEFLLENNPTLLAVDRAPMTYSIDPLGLASSDSFLQVEDPSN	120
QY	61	YXATXDSRLPGALVYDEFLLENNPTLLAVDRAPMTYSIDPLGLASSDSFLQVEDPSN	120

```

121  SSVRGAVNDLLAKKHODVGOVNNVPAKMOYEKITAHSMEOLKVKFGSPDEKGNLSDIDE 180
121  SSVRGAXXDLKWKHODVGOVNNVPAKMOYEKITAHSMEOLKVKFGSPDEKGNLSDIDE 180
181  NSVHSEKGIQIYNKQIYYTVDVAVKNGDVFQDTVEDLKORGISAERPLVYISXV 240
181  NSVHSEKGIQIYNKQIYYTVDVAVKNGDVFQDTVEDLKORGISAERPLVYISXV 240
241  AYKROYLLKLETTSSKDEVEAFALIKGVKAPQTEMKQIIDNTYKAVILGGPSSGA 300
241  AYKROYLLKLETTSSKDEVEAFALIKGVKAPQTEMKQIIDNTYKAVILGGPSSGA 300
301  RVTGKADWEDLIQSGRFTADHPLPISYTSFLRDVNVATFQNSTDYVEKTVAYRN 360
301  RVTGKADWEDLIQSGRFTADHPLPISYTSFLRDVNVATFQNSTDYVEKTVAYRN 360
361  GDLLDHSGAVYAQYYITWDELSDYHOGKEVLTTPKAMDNGODLTAHFTTSLPKGNVN 420
361  GDLLDHSGAVYAQYYITWDELSDYHOGKEVLTTPKAMDNGODLTAHFTTSLPKGNVN 420
421  LSVKIRECTGLAEMWMTVEKTDLPVRRKRTISWGTTLYPOVEDKVEN 471
421  LSVKIRECTGLAEMWMTVEKTDLPVRRKRTISWGTTLYPOVEDKVEN 471

```

```

RESULT 2
ENTRY 2
TITLE B43577 #type complete
ORGANISM perfringolysin O precursor - Clostridium perfringens
DATE 03-Feb-1993 #sequence_revision 03-Mar-1993 #text_change 13-Sep-1998

```

```

ACCESSIONS
REFERENCE
#authors Shimizu, T.; Okabe, A.; Minami, J.; Hayashi, H.
#journal Infect. Immun. (1991) 59:137-142
#title An upstream regulatory sequence stimulates expression of the perfringolysin O gene of Clostridium perfringens.
#cross-references M01D:9109951
#accession B43577
#molecule_type DNA
#residues 1-500 #label SHI
#cross-references GB:M01080; NID:9144885; PID:9144886
#note translation of the nucleotide sequence is not complete

```

```

REFERENCE
#authors Tweten, R.K.
#journal Infect. Immun. (1988) 56:3235-3240
#title Nucleotide sequence of the gene for perfringolysin O (theta-toxin) from Clostridium perfringens: significant homology with the genes for streptolysin O and pneumolysin.
#cross-references M01D:89032623
#accession A34951
#molecule_type DNA
#residues 1-125, 'EA', 129-500 #label TWE
#cross-references GB:M36704; NID:9144883; PID:9144884

```

```

REFERENCE
#authors Tweten, R.K.
#journal Infect. Immun. (1988) 56:3235-3240
#title Cloning and expression in Escherichia coli of the perfringolysin O (theta-toxin) gene from Clostridium perfringens and characterization of the gene product.
#cross-references M01D:89032622
#accession A60922
#molecule_type protein
#residues 29-45 #label TW2
#experimental_source ATCC 13124

```

```

GENETICS
#gene pfo; pfoA
#feature 1-28
#domain signal sequence #status predicted #label SIG
#product perfringolysin O #status experimental #label MAY
#length 500 #molecular-weight 55799 #checksum 4821

```

```

SUMMARY
#length 500 #molecular-weight 55799 #checksum 4821

```

```

Query Match 49.88; Score 1556; DB 2; Length 500;
Best Local Similarity 45.48; Pred. No. 3,556-265;
Matches 209; Conservative 99; Mismatches 151; Indels 1; Gaps 1;

```

```

35  NSIDSGISLSTNRNBNVLAASNDKIESFVPEKKGAKGKFTVEROKRSLTTSVPDISI 94
35  NSIDSGISLSTNRNBNVLAASNDKIESFVPEKKGAKGKFTVEROKRSLTTSVPDISI 94
3  NKAVNDELTMANNIDXXKLLTHQSESIENRFKXKGNOLPREFYXHEKKRSLSTNSDIXV 62
3  NKAVNDELTMANNIDXXKLLTHQSESIENRFKXKGNOLPREFYXHEKKRSLSTNSDIXV 62
95  IDSVNDRTPGALQDLADKAEVNRPTILWKRKPININIDPLGKENS-IVKDPPTIK 153
95  IDSVNDRTPGALQDLADKAEVNRPTILWKRKPININIDPLGKENS-IVKDPPTIK 153
63  XATXDSRLYPGALLVDEITXLENNPTILAVDRAPMTYSXXLGLASSDSFLOVEDPSNS 122
63  XATXDSRLYPGALLVDEITXLENNPTILAVDRAPMTYSXXLGLASSDSFLOVEDPSNS 122
154  VSGAIDELVSKNNEKXSTHTLPARTOYSSEMYSKSOSLALNVAKVLENSLGVFNA 213
154  VSGAIDELVSKNNEKXSTHTLPARTOYSSEMYSKSOSLALNVAKVLENSLGVFNA 213
123  VRGAXXDLKWKHODVGOVNNVPAKMOYEKITAHSMEOLKVKFGSPDEKGNLSDIDEFS 182
123  VRGAXXDLKWKHODVGOVNNVPAKMOYEKITAHSMEOLKVKFGSPDEKGNLSDIDEFS 182
214  VANNKRWMLIAYKQIFYTASADLPKNPSDLFDSYTFNDLKORGVSNAAPPLMNSVAY 273
214  VANNKRWMLIAYKQIFYTASADLPKNPSDLFDSYTFNDLKORGVSNAAPPLMNSVAY 273
183  VHSKEXKIQIYNKQIYYTVDVAVKNGDVFQDTVEDLKORGISAERPLVYISXVAY 242
183  VHSKEXKIQIYNKQIYYTVDVAVKNGDVFQDTVEDLKORGISAERPLVYISXVAY 242
274  GRTIYVLETTSSSKDVQAAPKALIKNTDIKNSQOKDIYENSFTAVYVIGDQAOENKV 333
274  GRTIYVLETTSSSKDVQAAPKALIKNTDIKNSQOKDIYENSFTAVYVIGDQAOENKV 333
243  XROYLLKLETTSSKDEVEAFALIKGVKAPQTEMKQIIDNTYKAVILGGPSSGAR 302
243  XROYLLKLETTSSKDEVEAFALIKGVKAPQTEMKQIIDNTYKAVILGGPSSGAR 302
334  VTKDFDELIRKVIKDNATFSTKPNAPYISYTSVFLKNSVAAVANKDYLRTSTESKGR 393
334  VTKDFDELIRKVIKDNATFSTKPNAPYISYTSVFLKNSVAAVANKDYLRTSTESKGR 393
303  VTGKADWEDLIQSGRFTADHPLPISYTSFLRDVNVATFQNSTDYVEKTVAYRNGD 362
303  VTGKADWEDLIQSGRFTADHPLPISYTSFLRDVNVATFQNSTDYVEKTVAYRNGD 362
394  INLDHSGAVYAQFVANDVSYDKEGNEVLTHKTWGSNYODKTAHSTVPLEANARIR 453
394  INLDHSGAVYAQFVANDVSYDKEGNEVLTHKTWGSNYODKTAHSTVPLEANARIR 453
363  LLDHSGAVYAQYYITWDELSDYHOGKEVLTTPKAMDNGODLTAHFTTSLPKGNVRLS 422
363  LLDHSGAVYAQYYITWDELSDYHOGKEVLTTPKAMDNGODLTAHFTTSLPKGNVRLS 422
454  IKARCTGLAEMWMTVEKTDLPVRRKRTISWGTTLYP 493
454  IKARCTGLAEMWMTVEKTDLPVRRKRTISWGTTLYP 493
423  VKIRECTGLAEMWMTVEKTDLPVRRKRTISWGTTLYP 462
423  VKIRECTGLAEMWMTVEKTDLPVRRKRTISWGTTLYP 462

```

```

RESULT 3
ENTRY 3
TITLE A43505 #type complete
ORGANISM listeriolysin O precursor - Listeria monocytogenes
DATE 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 17-Mar-1999

```

```

ACCESSIONS
REFERENCE
#authors Mengesha, J.; Vicente, M.F.; Chenevert, J.; Pereira, J.M.;
Geoffroy, C.; Gicquel-Sanzey, B.; Baquero, F.; Perez-Diaz,
J.C.; Cossart, P.
#journal Infect. Immun. (1988) 56:766-772
#title Expression in Escherichia coli and sequence analysis of the listeriolysin O determinant of Listeria monocytogenes.
#cross-references M01D:86153053
#accession A43505
#molecule_type DNA
#status preliminary
#residues 1-529 #label MEN
#cross-references GB:M24199; NID:9149652; PID:9149653
#note this sequence is derived from a strongly hemolytic strain, serotype 1/2c

```

```

REFERENCE
#authors Domann, E.; Chakraborty, T.
#journal Nucleic Acids Res. (1989) 17:6406
#title Nucleotide sequence of the listeriolysin gene from a Listeria monocytogenes serotype 1/2a strain.
#cross-references M01D:89366684
#accession S05306
#molecule_type DNA
#residues 1-529 #label DOM
#cross-references EMBL:X15127; NID:944106; PID:944107
#experimental_source strain EGD
#note this sequence is derived from a weakly hemolytic strain, serotype 1/2a

```

```

REFERENCE
A47606

```

[illegible]

ORGANISM	#formal_name	Listeria monocytogenes
DATE	#strain	120677
variety		12067
ACCESSIONS		22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 13-Sep-1998
REFERENCE		S24231
#authors		Rasmussen, O.F.; Beck, T.; Olsen, J.E.; Dons, L.; Rossen, L
#journal		Infect. Immun. (1991) 59:3945-9951
#title		Listeria monocytogenes isolates can be classified into two major types according to the sequence of the listeriolysin gene.
#cross-references	MUID:92040062	
#accession	S24231	
#status		nucleic acid sequence not shown; translation not shown
#molecule_type	DNA	
#residues	1-529	#label_RAS
#cross-references	EMBL:X60035; NID:g44110; PID:g44112	
#experimental_source	strain 12067, serotype 4b	
#note		the nucleotide sequence was submitted to the EMBL Data Library, June 1991
GENETICS		
#gene	lisa	
CLASSIFICATION	#superfamily	dipeptide transport protein
KEYWORDS		variance factor
FEATURE		
1-25		
26-529		
SUMMARY		
Query Match	46.9%;	Score 1468; DB 2; Length 529;
Best Local Similarity	40.6%;	Pred. No. 4.04e-248;
Matches	189;	Conservative 115; Mismatches 161; Indels 1; Gaps 1;
61	IDKYIIGDLNNNNLVYHGDAVNVPPRKQYKDGNEIYVEKKKSLNNNDIQVYNA	120
6	VNDFLANNYDXXKLLTHQGESIENRFKBEQNOLPXFVXXEKKRKSLSINTSDIAXYAT	65
121	ISSLTYPGALVKAANSELVENOPDVLVYKRSLSLSIDLPGMTNODNKIVYKNAITKSNVNN	180
66	XDSRLYPGALLVWETXLENNPRLTLADRAPMYKXXKLPGIASSDSFLQVEDPSNSVRC	125
181	AVNTLVSENNKRYEAOAYPNVSAKIDYDDEMAVSSQILAFGTFKFAVNNSLVNGEALIS	240
126	AXXDILAKHODYGOVN-NVPRKQYKXTAHSWEOQLKVFSGSEFKXGXSILIDFNFSVH	184
241	EGRKQEEVYSFKQIYVNVNNEPTPRSRFFGKAYATKEQLOALGVNAENPPAYISSVAYGR	300
185	SGEKKIQIYNNKKQIYVVSVAVNAKPNPGVQPDVTVEDLQNRGISARPLVYISVAYAKR	244
301	OYVLTLSNHSSTYKKAFFDAVSGKVSQDVELTNIKKSSFRKAVLYGGSARDEVOIID	360
245	OYVLTLETTSSKXEAFELINGVAVAPQTEKQILDTYKAVILIGDPPSGARVAY	304
361	GNLGDRLILKKGATFNERTPGVVIATYTFILKONELAAVIRKNNSEVETTSKAYTGGKIN	420
305	GKVDWVEDLIGESRFADHDPGLISTITTEFLRNVAATQNSTDIYETVYIYRNGDLI	364
421	IDHSGVYAQGNISWDELNDPEGNEIVQHKNNSENKSKLAHFTSSIYLPGNARINNY	480
365	LDHSGAVYAQYIITWXSLESDHOCKEVLTPKRAMDRNGODTLAFTTISIPLKGNVRLSVK	424
481	AKECGGLAMEMWRTVIDDRNLPLVKNNNISWGLTLYPKXNSVDN	526
425	IREGGLAMEMWRTVYEKTDLALYKRTISIWGTTILYPOVEDKVEN	470
RESULT	5	
ENTRY	S22341	#type complete
TITLE	ivanolysin precursor - Listeria ivanovii	
ORGANISM	#formal_name	Listeria ivanovii
DATE	07-Apr-1994	#sequence_revision 07-Apr-1994 #text_change 21-Aug-1998

```

ACCESSIONS      S22341; S36683
REFERENCE        S22340
#authors         Haas, A.; Dumbsky, M.; Kreft, J.
#journal         Blochim. Biophys. Acta (1992) 1130:81-84
#title           Listeriolysin genes: complete sequence of 110 from Listeria
                  ivanovii and of 1so from Listeria seeligeri.
#cross-references MUID:92182018
#accession       S22341
##molecule_type DNA
##residues       1-528 ##label HAS
##cross-references EMBL:X60461
#note            the authors translated the codon ACA for residue 331 as
                  Val

REFERENCE        S36683
#authors         Kreft, J.
#submission      Submitted to the EMBL Data Library, July 1991
#accession       S36683
##molecule_type DNA
##residues       1-319,'T',321-528 ##label KRE
##cross-references EMBL:X60461

GENETICS
#gene            110
CLASSIFICATION . #superfamily dipeptide transport protein
FEATURE
1-24             #domain signal sequence #status predicted #label SIG\
25-528          #product lvanolysin #status predicted #label MAT
SUMMARY          #length 528 #molecular-weight 58511 #checksum 6874

Query Match      46.7% Score 1460; DB 2; Length 528;
Match Local Similarity 40.8%; Pred. No.1,43e-246;
Matches 190; Conservative 116; Mismatches 159; Indels 1; Gaps 1

Db 60 IDQYIGLDYDKNNILVYDGEAVKKNVPKAGYKEGNOYIVVEKKKKSINONNADIQVINS 119
   V::F::I::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 6 VNDFTIAMNYDXKKLLTHQGESIENRFRKEGNQLPXEFEVXXEKKRSISTNTSDIXAYAT 65
   ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 120 LASLTYPGALYKANSELVENQPDVLPEKRDSVTLSIDLPGWVNHNETLVGNMTRKSNIND 179
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Oy 66 XDSRLTPGALLVADDTXLENNPTLLVADPAKPITSXLPLGLASSDSPSLQVEDPSRSSVRG 125
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Db 180 GVNLTVDNRNNKYSEELYENISAKIDYIDDEMAYESOLVAKTGAFAKAVNSLNWNFGAIS 239
   I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 126 AXDDLAKWHQDDGVQN-NVPRAXQYEKKXTAHSMQLKVKGSGDPEKKXSNDIDFNFSVH 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Oy 240 ESKVOEDVINFEQIIYYTVANNEPTSPSRFFGKSYTKENLQALGVAAENPPRAYISSVAAGR 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 185 SGEXKIQQVANKQIIYYTVSVDAVKKPGDFDPTIVEDLKORGISAERPPLYISXVAIYR 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 300 DIFVKLSTSSHSTRKAADFARFKKSVKGDTELENIIONASFKAVIYGSAKDVEELIID 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 245 QYLKLETFTSSKXEVBAAFELIKGVYKAPQEMKQIIDNTYKVAIVLIGDPSSGARVYT 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Oy 360 GDLSLRDLILKCGANDPKKNPCPVLAITYTNFLKDNQALVYKKNSEYIEFTSKAYSDBGIN 419
   I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 305 GYDVWDIEDIQGCSRTDHPGLPISTYSFLRDVVNAAFONSTDYEVETKYATYANGDLL 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 420 LPHSGLYVRFNVWDEVDEYDANGNGENVVHKMKSENCKKLAEHTSIYLPQNAENINH 479
   I H S G A Y V R F N W D E V D E Y D A N G N G E N V V H K M K S E N C K K L A E H T S I Y L P Q N A E N I N H
Oy 365 LDHSGAYVAQYIITWKXELSIDHQGRFVLTPKAMDRNGDULTAHTFTSIPDKGNVNLISVK 424
   I D H S G A Y V A Q Y I I T W K X E L S I D H Q G R F V L T P K A M D R N G D U L T A H T F T S I P D K G N V N L I S V K

Db 480 AECTGLAEMWRIVVDRNRLPLVKNRNVCIMGTLTYPYASDTVDN 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 425 IRECTGLAEMWRIVYEKTDLXIVRRRTISIGTTLTPQVEDKVEN 470
   I R E C T G L A E M W R I V Y E K T D L X I V R R R T I S I G T T L T P Q V E D K V E N

RESULT 6
ENTRY    A37858 #type complete
TITLE    alveolysin - Bacillus alvei
ORGANISM #formal_name Bacillus alvei
DATE     28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change
          09-Sep-1997
ACCESSIONS A37858
REFERENCE   A37858

```

Accessions	Reference	Annotations	Sequence	Score	DB	Length	Indels	Gaps
1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18
20	21	22	23	24	25	26	27	28
30	31	32	33	34	35	36	37	38
40	41	42	43	44	45	46	47	48
50	51	52	53	54	55	56	57	58
60	61	62	63	64	65	66	67	68
70	71	72	73	74	75	76	77	78
80	81	82	83	84	85	86	87	88
90	91	92	93	94	95	96	97	98
100	101	102	103	104	105	106	107	108
110	111	112	113	114	115	116	117	118
120	121	122	123	124	125	126	127	128
130	131	132	133	134	135	136	137	138
140	141	142	143	144	145	146	147	148
150	151	152	153	154	155	156	157	158
160	161	162	163	164	165	166	167	168
170	171	172	173	174	175	176	177	178
180	181	182	183	184	185	186	187	188
190	191	192	193	194	195	196	197	198
200	201	202	203	204	205	206	207	208
210	211	212	213	214	215	216	217	218
220	221	222	223	224	225	226	227	228
230	231	232	233	234	235	236	237	238
240	241	242	243	244	245	246	247	248
250	251	252	253	254	255	256	257	258
260	261	262	263	264	265	266	267	268
270	271	272	273	274	275	276	277	278
280	281	282	283	284	285	286	287	288
290	291	292	293	294	295	296	297	298
300	301	302	303	304	305	306	307	308
310	311	312	313	314	315	316	317	318
320	321	322	323	324	325	326	327	328
330	331	332	333	334	335	336	337	338
340	341	342	343	344	345	346	347	348
350	351	352	353	354	355	356	357	358
360	361	362	363	364	365	366	367	368
370	371	372	373	374	375	376	377	378
380	381	382	383	384	385	386	387	388
390	391	392	393	394	395	396	397	398
400	401	402	403	404	405	406	407	408
410	411	412	413	414	415	416	417	418
420	421	422	423	424	425	426	427	428
430	431	432	433	434	435	436	437	438
440	441	442	443	444	445	446	447	448
450	451	452	453	454	455	456	457	458
460	461	462	463	464	465			

Db 228 AIDNLVNMQHDNYSGGNTLPARTQYTESMWSKQIEALNVNSKILDTGLIDPKSISK 2878

[illegible]

Db 403 LDHGAAYAOEDVSWDGFEDQNGKEILLHKTWEGSGDKTAHASTVLPSPNSKNIKIV 462
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 365 LDHSAIYAQYIITWKEISYDHQGEVLPRKMDNGDILAHFTTSLPLGNVRLNSVK 424
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 463 ARECTGLAWEMWRTI 477
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 425 IRECTGLAWEMWRTI 439
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
RESULT 10
ENTRY S47298 #type complete
TITLE sul1ysin - Streptococcus suis
ORGANISM #formal_name Streptococcus suis
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
ACCESSIONS S47298
REFERENCE S47297
#authors Segers, R.
#submission submitted to the EMBL Data Library, September 1994
#accession S47298
#status Preliminary
#molecule_type DNA
#residues 1-112 #label SEG
SUMMARY ##cross-references EMBL:236908; NID:g535307; PID:g535308
#length 112 #molecular_weight 12464 #checksum 9264
Query Match 14.8%; Score 464; DB 2; Length 112;
Best local Similarity 50.9%; Pred. No. 3.51e-58;
Matches 57; Conservative 22; Mismatches 33; Indels 0; Gaps 0;
Db 1 PGVPSVSTTFVXXXXXPAOXLSNSETTSTVHNSALTLXSGAYVAKYITNGEVS 60
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 325 PGPISTTSTFSLDNVAFVAFONSTDVEETKVTAYRNGDILLDHSGAYVAYIITWKEIS 384
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 61 NEAGEVWEKAMDKNGVNLTSWSETIXIPGNAXNLHYNIOECTGLAWEMW 112
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 385 DHQGEVLPRKMDNGDILAHFTTSLPLGNVRLNSVKIRECTGLAWEMW 436
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
RESULT 11
ENTRY S47297 #type complete
TITLE sul1ysin - Streptococcus suis
ORGANISM #formal_name Streptococcus suis
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
ACCESSIONS S47297
REFERENCE S47297
#authors Segers, R.
#submission submitted to the EMBL Data Library, September 1994
#accession S47297
#status Preliminary
#molecule_type DNA
#residues 1-96 #label SEG
SUMMARY ##cross-references EMBL:236907; NID:g535305; PID:g535306
#length 96 #molecular_weight 10471 #checksum 6662
Query Match 6.7%; Score 209; DB 2; Length 96;
Best local Similarity 32.6%; Pred. No. 1.01e-14;
Matches 29; Conservative 22; Mismatches 38; Indels 0; Gaps 0;
Db 1 KQONQYQSILTYGQELTNEGEVINDPPTTGMLENGRVRVLRREKXITITNSADITVI 60
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 4 KAVNDFLIANNYXXKLLTHQGESIENREKESNQLPXEYXXERRKSLSTVTSIYX 63
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 61 XAKANISPGALLRAXONLDDNNPTLISI 89
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 64 ATXDSRLYPGALLVDETXLENNPTLLAV 92
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
RESULT 12
ENTRY S60947 #type complete
TITLE hypothetical protein YOR220w - yeast (Saccharomyces cerevisiae)

ALTERNATE_NAMES
ORGANISM hypothetical protein O5050; hypothetical protein YOR50-10
#formal_name Saccharomyces cerevisiae
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 05-Jun-1998
ACCESSIONS S60947; S67113; S71722
REFERENCE S60938
#authors Gailson, F.; Dujon, B.
#submission submitted to the EMBL Data Library, October 1995
#description Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of the yeast Saccharomyces cerevisiae.
#accession S60947
#molecule_type DNA
#residues 1-265 #label GAL
SUMMARY ##cross-references EMBL:X92441; NID:g1050762; PID:g1050772
S67104
REFERENCE S67104
#authors Boyer, J.; Fairhead, C.; Gallon, L.; Gailson, F.; Michaux, G.; Thierry, A.; Dujon, B.
#submission submitted to the Protein Sequence Database, July 1996
#accession S67113
#molecule_type DNA
#residues 1-265 #label BOY
SUMMARY ##cross-references EMBL:275128; NID:g1420509; PID:e252085; PID:g1420510; MIPS:YOR220w
S71713
REFERENCE S71713
#authors Gailson, F.; Dujon, B.
#journal Yeast (1996) 12:877-885
#title Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of the yeast Saccharomyces cerevisiae.
#cross-references MUID:9643797
#accession S71722
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-265 #label GAW
SUMMARY ##cross-references EMBL:X92441; NID:g1050762; PID:g1050772
#note the nucleotide sequence was submitted to the EMBL Data Library, October 1995
GENETICS
map_position 15R
SUMMARY #map_position 15R
#length 265 #molecular_weight 29255 #checksum 6028
Query Match 3.6%; Score 113; DB 2; Length 265;
Best local Similarity 35.2%; Pred. No. 2.77e-01;
Matches 25; Conservative 10; Mismatches 33; Indels 3; Gaps 3;
Db 118 LLEDHRNKPPLISINT-DPGVTGVDSLSLNGKSSLSPEKSSLESPMLKLSIDSKPFSIQ 176
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 43 FVXXERRKRSLSNTSDIXVATXDSRLY-PCALLVDETXLENNPTL-LAVDRAPMRTS 100
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 177 EPLPKLSRSSS 187
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 101 XPLGLASSDS 111
||| ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
RESULT 13
ENTRY S03974 #type complete
TITLE amine oxidase (flavin-containing) (EC 1.4.3.4) A - bovine
ALTERNATE_NAMES #formal_name Bos primigenius taurus #common_name cattle
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 29-Jan-1999
ACCESSIONS S03974
REFERENCE S03974
#authors Powell, J.F.; Hsu, Y.P.; Weyler, W.; Chen, S.; Salach, J.; Andrikopoulos, K.; Maillet, J.; Breakfield, X.O.
#journal Biochem. J. (1989) 259:407-413
#title The primary structure of bovine monamine oxidase type A. Comparison with peptide sequences of bovine monamine oxidase type B and other flavoenzymes.
#cross-references MUID:89246344
#accession S03974
#status not compared with conceptual translation
#molecule_type mRNA

```
##residues 1-527 ##label POW
##cross-references GB:X15609; NID:g2523; PID:g2524
KEYWORDS FAD; flavoprotein; mitochondrion; oxidoreductase
FEATURE 15-43 #region beta-alpha-beta FAD nucleotide-binding fold
406 #modified_site S-(8alpha-FAD)-cysteine (Cys) #status
#predicted
SUMMARY #length 527 #molecular-weight 59800 #checksum 6598

Query Match 3.6%; Score 112; DB 2; Length 527;
Best Local Similarity 18.7%; Pred. No. 3.63e-01;
Matches 28; Conservative 42; Mismatches 74; Indels 6; Gaps 5;

Db 1 MESLQKTSAGQMDVNVIGGIGSLAAK-LLAHEENVLVLEAREVGGRTYVRNH 59
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 250 LETSXSXEVEAFALIKGVAPQTEWKQILDNTXKAVILIGDPSGARVYTKVDM 309
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 60 V-DYVDVGGAYVPTQNRILRLSKQLGLEYKVNVERLVHYKGTYPFGAFPPVNP 118
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 310 VEDLIEGSRFTDHPQLPISYTTSF-LRDNVVATFQNSTDYETKVTAYRNGDLLDS 368
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 119 IAYL-DYNNILMRTM--DNMGKEIPADAPWE 145
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 GAYVAQYITTWXELSTDHCKEVLTPKAMD 398
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
ENTRY A70385 #type complete
TITLE DNA gyrase A subunit Aquifex aeolicus
ORGANISM #formal_name Aquifex aeolicus
DATE 08-May-1998 #sequence_revision 08-May-1998 #text_change
16-Dec-1998
ACCESSIONS A70385
REFERENCE A70300
#authors Deckert, G.; Warren, P. V.; Gaasterland, T.; Young, W. G.;
Lenox, A. L.; Graham, D. E.; Overbeek, R.; Sneed, M. A.;
Keller, M.; Aujay, M.; Huber, R.; Feldman, R. A.; Short,
J. M.; Olson, G. J.; Swanson, R. V.
#journal Nature (1998) 392:353-358
#title The complete genome of the hyperthermophilic bacterium
Aquifex aeolicus.
#cross-references NUID:98196666
#accession A70385
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-744 #label AOF
##cross-references GB:AE000716; NID:g2983478; PID:g2983485; GB:AE000657
#experimental_source strain VFS
GENETICS
#gene gyrA
CLASSIFICATION #superfamily DNA topoisomerase (ATP-hydrolyzing) chain A;
phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain
homology
T77777RE
4-237 #domain phage T4 DNA topoisomerase (ATP-hydrolyzing)
medium chain homology #label TOP
SUMMARY #length 744 #molecular-weight 84080 #checksum 7466

Query Match 3.6%; Score 112; DB 2; Length 744;
Best Local Similarity 24.2%; Pred. No. 3.63e-01;
Matches 22; Conservative 26; Mismatches 37; Indels 6; Gaps 5;

Db 449 IEY-YKKLVASEGERI-KVFIETELVKKY--GD-KRRTFIGVKEVKGSTIVAVLD 503
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 13 MNDYXXKLTHQGESIENRFKXEGNLPKXEFYXXKKRSLSTNTSDIXVXAFDRLYP 72
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 504 GSTIPEELPLEKAP-VNILRVPTFEGDLF 533
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 73 GALLVDETLENNPTLLAVDRAPMTYSXXL 103
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
```

```
ENTRY T03415 #type complete
TITLE S-layer protein - Campylobacter rectus
ORGANISM #formal_name Campylobacter rectus
DATE 16-Mar-1999 #sequence_revision 16-Mar-1999 #text_change
16-Mar-1999
ACCESSIONS T03415
REFERENCE Z14942
#authors Wang, B.; Kraig, E.; Kolodrubetz, D.
#journal Infect. Immun. (1998) 66:1521-1526
#title A new member of the S-layer protein family: characterization
of the crs gene from Campylobacter rectus.
#accession T03415
#status preliminary
##residues 1-1361 ##label WAN
##cross-references EMBL:AF010143; NID:g2459960; PID:g2459961;
PID:g2459961
SUMMARY #length 1361 #molecular-weight 144385 #checksum 6731

Query Match 3.6%; Score 114; DB 3; Length 1361;
Best Local Similarity 19.9%; Pred. No. 2.11e-01;
Matches 37; Conservative 50; Mismatches 89; Indels 10; Gaps 9;

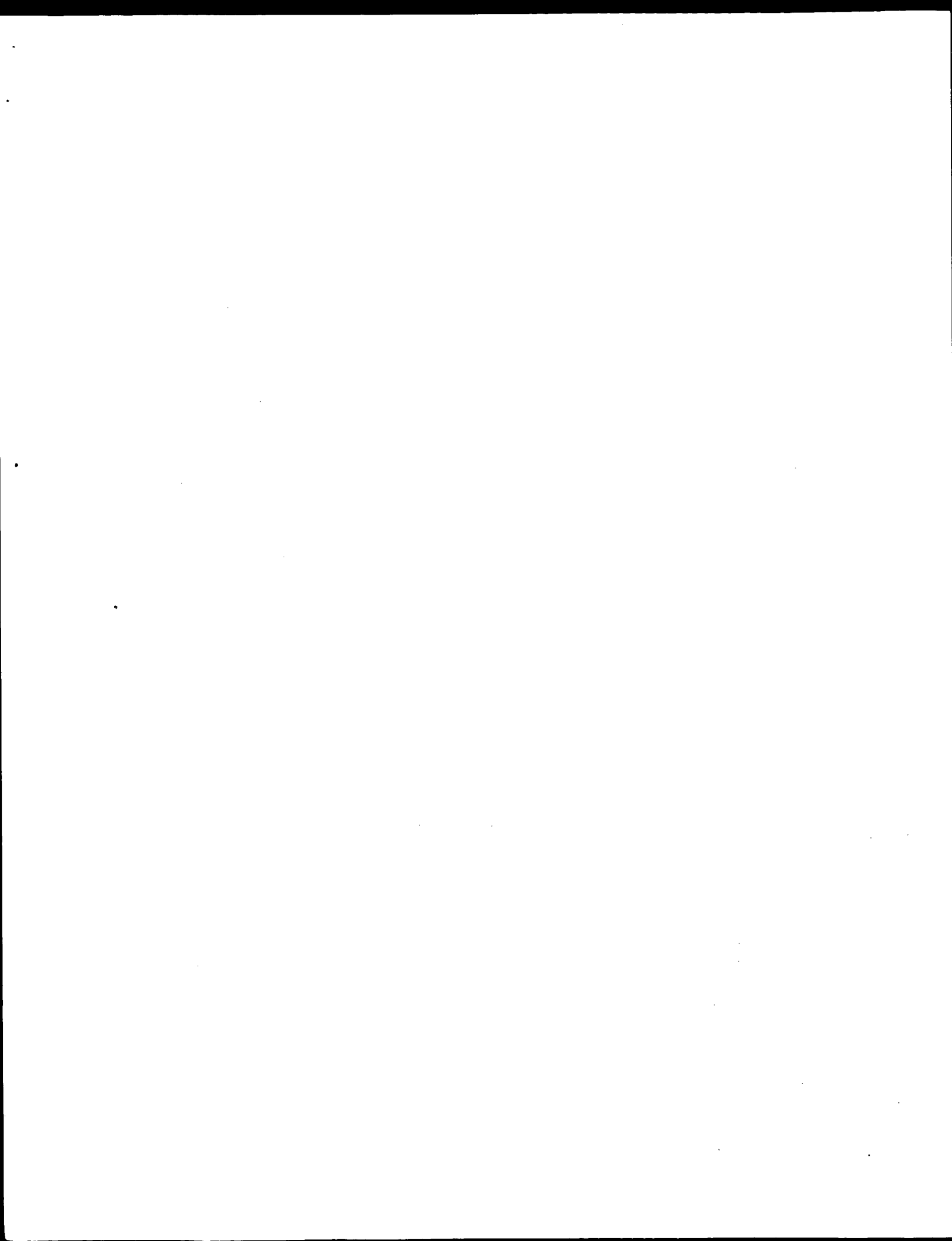
Db 178 VDMHSNPEHPGKAYELTNTDNTATNFNAPMK-HNPGTDRIWTLQSSDKLTGDSRHD 236
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 115 VEDPSNSSVRCAXDXDLAKMWDGYGVNVPARXQYKXTAHSNE-OLKYKFGSDEKKG 173
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 237 NTLNVEFGQANADGDPSTRTPTLTNTQININIEVTGTNTLDL-RSDNDYEKINIRITK 295
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 NSLDIDFNSVHS--GEXXIQIVNKKQI--YTVSVDAVKNPGDVQDITVEDLKORISIA 230
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 296 EAGKKFNVESTG-OKLVGMRLANAK-KDIDYKFEH-KKGVLSGFEDKSNVFLDNVAKS 352
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 231 ERPLVY-ISKVAYIKROYLLKLETTSSXSXEVAEALIKGVKVAQTEWKQILDNTXVKA 289
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 353 LSTSD 358
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 VILGSD 295
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: Mon Aug 30 16:13:46 1999
Job time : 52 secs.
```



 WIRE (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Search: 77977 seqs, 28268293 residues
 Listing first 45 summaries
 Database: swiss-prot37
 1:swissprot
 Statistics: Mean 51.211; Variance 95.336; scale 0.537

Title: >US-09-120-044-4
 Description: (1-471) from US09120044.pep
 Perfect Score: 3127
 Sequence: 1 MARKAVNDPILANNVXXKL.....TISINGTITYPEVEKQVEND 471

Scoring table: PAM 150
 Gap 11

Post-processing: Minimum Match 0%

Database: swiss-prot37
 1:swissprot

Statistics: Mean 51.211; Variance 95.336; scale 0.537

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	3116	99.6	470	1	TACY_STRPN	0.00e+00
2	1556	49.8	500	1	TACY_CLOPE	0.00e+00
3	1475	47.2	529	1	TACY_LISMO	2.10e-290
4	1458	46.6	528	1	TACY_LISTIV	1.43e-286
5	1416	45.3	501	1	TACY_PAEAL	4.14e-277
6	1414	45.2	530	1	TACY_LISEE	1.17e-276
7	1379	44.1	571	1	TACY_STRPO	8.80e-269
8	1380	44.1	574	1	TACY_STRPO	5.24e-269
9	1363	43.6	574	1	TACY_STRPO	3.49e-265
10	1298	41.5	485	1	TACY_STRPO	1.40e-250
11	1112	3.6	527	1	HOFA_BOVIN	6.96e-02
12	108	3.5	462	1	MEPB_YEAST	2.34e-01
13	108	3.5	1228	1	SLAP_BACST	2.34e-01
14	107	3.4	766	1	GCH2_POLE	3.15e-01
15	105	3.4	374	1	ARCB_ECOLI	5.67e-01
16	105	3.4	780	1	CTPA_MYCLE	1.35e+00
17	102	3.3	514	1	THDI_SALTY	1.01e+00
18	103	3.3	722	1	THDI_SALTY	1.01e+00
19	102	3.3	878	1	SYV_METTA	1.35e+00
20	99	3.2	279	1	YRR6_MYCCA	3.14e+00
21	101	3.2	350	1	ARGI_SOYBN	1.79e+00
22	100	3.2	427	1	GLYA_SYMY3	2.38e+00
23	100	3.2	502	1	SPEI_LYCES	2.38e+00

Result ID	Score	Match	Length	ID	Description	Pred. No.
24	100	3.2	511	1	NOPS_YEAST	2.38e+00
25	100	3.2	887	1	MTP_BOVIN	2.38e+00
26	101	3.2	1026	1	BGAL_STRTR	1.79e+00
27	100	3.2	1123	1	V120_HSV11	2.38e+00
28	99	3.2	3027	1	POLG_PIV1	3.14e+00
29	98	3.1	235	1	KEEP_CSV	4.15e+00
30	97	3.1	236	1	DEGU_BACBR	5.46e+00
31	97	3.1	325	1	SBCD_BACPO	5.46e+00
32	98	3.1	363	1	RIBB_PHOPO	5.46e+00
33	98	3.1	418	1	Y285_HELPY	4.15e+00
34	97	3.1	419	1	Y223_MTCPN	5.46e+00
35	96	3.1	457	1	FKB4_RABIT	5.46e+00
36	97	3.1	486	1	YDBH_SCHPO	5.46e+00
37	98	3.1	513	1	YJIV_ECOLI	5.46e+00
38	96	3.1	533	1	NADP_SYNY3	4.15e+00
39	97	3.1	630	1	PRLR_ORENT	7.15e+00
40	98	3.1	743	1	BGAL_THET	5.46e+00
41	96	3.1	919	1	RPO2_CAPVK	7.15e+00
42	96	3.1	1076	1	VIA_CMY11	7.15e+00
43	97	3.1	1309	1	HSEK_CAVPO	5.46e+00
44	96	3.1	1309	1	CEK1_SCHPO	7.15e+00
45	98	3.1	1808	1	TENA_CHICK	4.15e+00

ALIGNMENTS

RESULT 1
 ID TACY_STRPN STANDARD: PRT: 470 AA.
 AC P11990;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
 DT 15-NOV-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE PNEUMOLYSIN (THIOL-ACTIVATED CYTOLYSIN).
 GN PLY.
 OS STREPTOCOCCUS PNEUMONIAE.
 OC BACTERIA: FIRMICUTES: BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
 OC STREPTOCOCCUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NCTC 7466 SEROTYPE II;
 RX MEDLINE: 87193109.
 RA WALKER J.A., ALLEN R.L., FLAMAGNE P., JOHNSON M.K., BOUINOS G.J.;
 RT "Molecular cloning, characterization, and complete nucleotide
 sequence of the gene for pneumolysin, the sulfhydryl-activated toxin
 of Streptococcus pneumoniae."
 RL INFECT. IMMUN. 55:1184-1189(1987).
 CC -!- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
 CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
 EUKARYOTIC CELL MEMBRANES.
 CC -!- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: X52474; GA7404; -
 CC EMBL: M17717; G153692; -
 CC PIR: A28568; A28568;
 CC PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
 CC PFAM: PF01289; Thiol_cytolysin; 1.
 CC HSSP: P19995; TPFO.
 CC TOXIN: HEMOLYSIN; CYTOLYSIS.
 CC INT_MET 0
 CC ACT_SITE 427 427 BINDING TO CHOLESTEROL (BY SIMILARITY).
 CC SEQUENCE 470 AA; 52768 MW; D3F3A252 CRC32;
 Query Match 99.6%; Score 3116; DB 1; Length 470;
 Best Local Similarity 94.5%; Pred. No. 0.00e+00;

	Matches	144;	Conservative	0;	Mismatches	26;	Indels	0;	Gaps	0;	
Db	1	ANKAVNDPII	AMNTDKKKLLTHG	SESIENPR	IKGNGNLP	DEFFVYIE	KKRSLS	NTSDIS	60		
QY	2	ANKAVNDPII	AMANDYKKLLTHG	SESIENPR	KKGKGNLP	EEFYVX	KKRRSL	NTSDIS	61		
Db	61	VTANRDSKLY	PGALLVYDETL	ENNNP	TLAVDRA	PMTYSIDLP	GLASSD	SFLQVEDPS	120		
QY	62	VXAIYXDSLY	PGALLVYDETL	ENNNP	TLAVDRA	PMTYSIXXLP	GLASSD	SFLQVEDPS	121		
Db	121	SVRGAVNDL	LAKKHODY	GVQVNNV	PARMOY	EKITAH	SMQLKVK	SGDSFE	KSGNSL	180	
QY	122	SVRGXDXDL	LAKKHODY	GVQVNNV	PARMOY	EXXIAH	SMQLKVK	SGDSFE	KSGNSL	181	
Db	181	SVHSGEKO	IOAVNF	KOIIYTV	SVDAVK	NRGQV	FODTV	VEDL	KORGIS	AEKRP	240
QY	182	SVHSGEKK	IOIYNN	KOIIYTV	SVDAVK	NRGQV	FODTV	VEDL	KORGIS	AEKRP	241
Db	241	YGRQVY	LLKLETT	SKSDE	VEAAR	EALIK	GVAP	APOT	EWKOL	INDTE	300
QY	242	YXROVY	LLKLETT	SKSXE	VEAAR	EALIK	GVAP	APOT	EWKOL	INDTE	301
Db	301	VYTGVDV	WVEDL	IOGSR	FP	TADH	PLP	SYTTS	FLRD	NVVA	360
QY	302	VYTGVDV	WVEDL	IOGSR	FP	TADH	PLP	SYTTS	FLRD	NVVA	361
Db	361	DLILDH	SGAAYA	QOYYIT	WDEL	SYDH	OG	EVIL	PKAMD	NGOD	420
QY	362	DLILDH	SGAAYA	QOYYIT	WDEL	SYDH	OG	EVIL	PKAMD	NGOD	421
Db	421	SVKIRE	CGI	AMEM	RTY	EKTD	ILV	KRT	SI	AGT	470
QY	422	SVKIRE	CGI	AMEM	RTY	EKTD	ILV	KRT	SI	AGT	471

RESULT	2	STANDARD;	PRT;	500 AA.
ID	TACY CLOPE			
AC	P19995:			
DT	01-FEB-1991 (REL. 17, CREATED)			
DT	01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)			
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)			
DE	PERRINGOLYSIN O PRECURSOR (THETA-TOXIN) (THIOL-ACTIVATED CYTOLYSIN).			
GN	PFO OR PFOR OR PFOA.			
OS	CLOSTRIDIUM PERRINGENS.			
OC	PLASMID PTB1B.			
OC	BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;			
RN	CLOSTRIDIUM.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 9109951.			
RA	SHIMIZU T., OKABE A., MINAMI J., HAYASHI H.;			
RT	"An upstream regulatory sequence stimulates expression of the			
RT	perringolysin O gene of <i>Clostridium perringsens</i> .";			
RL	INFECT. IMMUN. 59:137-142(1991).			
RL	[2]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 29-44.			
RC	STRAIN-ATCC 13124;			
RX	MEDLINE: 89032623.			
RA	TWETEN R.K.;			
RT	"Nucleotide sequence of the gene for perringolysin O (theta-toxin)			
RT	from <i>Clostridium perringsens</i> : significant homology with the genes for			
RT	streptolysin O and pneumolysin.";			
RL	INFECT. IMMUN. 56:3235-3240(1988).			
RN	[3]			
RP	SEQUENCE OF 29-45 AND 305-312.			
RC	STRAIN-PB6K;			
RX	MEDLINE: 87076517.			
RA	OHNO-IWASHITA Y., IWAMOTO M., MITSUI K., KAWASAKI H., ANDO S.;			
RT	"Cold-labile hemolysin produced by limited proteolysis of theta-toxin			
RT	from <i>Clostridium perringsens</i> .";			
RL	BIOCHEMISTRY 25:6048-6053(1986).			
RN	[4]			
RP	SEQUENCE OF 492-500 FROM N.A.			

```

RN STRAIN-NCTC 8237;
RX MEDLINE: 96123363.
RA SHIMIZU T., KOBAYASHI T., BA-THEIN W., OHTANI K., HAYASHI H.;
RT "Sequence analysis of flanking regions of the pta gene of
RT Clostridium perfringens; beta galactosidase gene (pbq) is located in
RT the 3'-flanking region."
RL MICROBIOL. IMMUNOL. 39:677-686(1995).
RN [5]
RP IMPORTANCE OF THIOL-GROUP IN CHOLESTEROL BINDING.
RX MEDLINE: 88004463.
RA IWAMOTO M., OHNO-TWASHITA Y., ANDO S.;
RT "Role of the essential thiol group in the thiol-activated cytolysin
RT from Clostridium perfringens."
RL EUR. J. BIOCHEM. 167:425-430(1987).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE: 97113505.
RA FELT S.C., ROSSJOHN J., ROHDE K., TWETEN R.K., PARKER M.W.;
RT "Crystallization and preliminary x-ray analysis of a thiol-activated
RT cytolysin."
RL FEBS LETT. 397:290-292(1996).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE: 97325744.
RA ROSSJOHN J., FELT S.C., MCKINSTRY W.J., TWETEN R.K., PARKER M.W.;
RT "Structure of a cholesterol-binding, thiol-activated cytolysin and a
RT model of its membrane form."
RL CELL 89:685-692(1997).
CC -I- FUNCTION: SULPHYRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -I- SUBUNIT: FORMS OLIGOMERS IN THE HOST MEMBRANE.
CC -I- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M81080; G144886; -
DR EMBL, M36704; G144884; -
DR EMBL, D49537; G1502275; -
DR PIR, B43577; B43577.
DR PDB: 1PFO; 05-AUG-98.
DR PROSITE, PS00481; THIOL_CYTOLYSIN; 1.
DR PFAM: PF01289; Thiol_cytolysin; 1.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMID; 3D-STRUCTURE.
FT SIGNAL 1 28
FT CHAIN 29 500 PERFRINGOLYSIN O.
FT ACT SITE 459 459 BINDING TO CHOLESTEROL.
FT CONFLICT 30 30 D -> K (IN REF. 3).
FT CONFLICT 34 34 K -> I (IN REF. 3).
FT CONFLICT 126 128 RRP -> EA (IN REF. 2).
SQ SEQUENCE 500 AA; 55799 MW; 17FC32BC CRC32;
Query Match 49.8%; Score 1556; DB 1; Length 500;
Best Local Similarity 45.4%; Pred. No.0.00e+00;
Matches 209; Conservative 99; Mismatches 151; Indels 1; Gaps 1;
Db 35 NOSIDISGSISYRNVEYSLASNGDKIESFPKBEKAGNKFIVERORRSLLTSPVDISI 94
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 3 NKAAYVDFLANVYXXKLLTHQGESIENRXXKGNQLPXFEXXERKKRSISTMTSDIXV 62
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 95 IDSVNDRYPCGALADAKAYENRPTILMYKKRPITINIDPLGLGGENS-IKYVDPTYK 153
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 63 XATXDSRYLPKGLLVVETXLENNPTLLADRAPMYYSKXLLGGLASDSFLOVEDPSNSS 122
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 154 VSGAIDELVSKWNEKYSHTLRLPRTQYSSEWYSKSOISSLANVAKYLENSLGVDFNA 213
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

QY 123 VRAXXDLAKMHODYGOVNVNPARXQYEXKTAHSEQLKVFSGPSEKXNSLIDFNS 182
DB 214 VANNEKVMILAYKQJFYTSADLPKRNPSDLPDDSYTFNDLKKGVSNAPPLAMSNVAY 273
QY 183 VHSGEKXIQVNNKQIYITVSDAVKNGVDYIVEDLQKRGISARPLVYISXVAY 242
DB 274 GRITVYKLETTSSSKDVQAFAKALIKNTDIKNSQYKDIYENSFPAAVVLGGDAOEHNAY 333
QY 243 XROVYIKLETTSSXSEVEAFALIKGVAPQTEWKQILNDNTYKAAVILGGDPSSGARV 302
DB 334 VTKDFEIRKIVKDNATFEKRNAPYISTSVFLKNSVAAYNKTDYLETSTESKCK 393
QY 303 VTKGVVMEDELIOGSRFTADHPGLPISTYTFSLRDNVAVATFONSDYVEIKVATVRNGD 362
DB 394 INLDHSGAVYAFVAMDEVSADKEGNEVLTFRKTMGDNQODKTAHSTYIPEANRNR 453
QY 363 LLLDHSGAVYAOYITVWELSLDHOKEVLTFRKMRNODDLTAHFTYISPLKGVNRNL 422
DB 454 IKARECTGLAMEMWMDYISEYDPLTNNINVSIGTTLTP 493
QY 423 VKIRECTGLAMEMWRTVYEKTDLXVKKRTISIMGTTLTP 462

RESULT 3
ID TACY LISMO STANDARD. PRT: 529 AA.
AC P13128: Q48747; Q57096; Q57206;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LISTERIOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN HLY OR LISA OR HLYA
OS LISTERIA MONOCYTOGENES.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC LISTERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EGD / SEROTYPE 1/2A.
RX MEDLINE: 89366684.
RA DOMANN E., CHARRABORTY T.;
RT "Nucleotide sequence of the listeriolysin gene from a Listeria
RL monocytogenes serotype 1/2a strain.";
RN NUCLEIC ACIDS RES. 17:6406-6406(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88153053.
RA MENGAUD J., VICENTE M.-F., CHENEVERT J., PEREIRA J.M., GEFEROY C.,
RT "Expression in Escherichia coli and sequence analysis of the
RL listeriolysin O determinant of Listeria monocytogenes.";
RN INECT. IMMUN. 56:766-772(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-12067;
RX MEDLINE: 92040062.
RA RASMUSSEN O.F., BECK T., OLSEN J.E., DONS L., ROSSEN L.;
RT "Listeria monocytogenes isolates can be classified into two major
RL types according to the sequence of the listeriolysin gene.";
RN INECT. IMMUN. 59:3943-3951(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-F2365, F5782 (SEROTYPE 4B), AND F4233, F6789 (SEROTYPE 1/2B);
RA VINES A., SWAMINATHAN B.;
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15127; G44107; -
DR EMBL: M24199; G149653; -
DR EMBL: X60035; G44112; -
DR EMBL: U25452; G887870; -
DR EMBL: U25443; G887016; -
DR EMBL: U25446; G887028; ALT_INIT.
DR EMBL: U25449; G887864; ALT_INIT.
DR PIR: S05306; S05306.
DR PIR: A43505; A43505.
DR PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM: PF01289; Thiol_cytolysin; 1.
DR HSP: P19995; IPEO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 529
FT ACT_SITE 484 484
FT VARIANT 35 35
FT VARIANT 39 39
FT VARIANT 438 438
FT VARIANT 523 523
SQ SEQUENCE 529 AA; 5868 MW; 269EA737 CRC32;

Query Match 47.2%; Score 1475; DB 1; Length 529;
Best local Similarity 40.8%; Pred. No. 2,10e-290;
Matches 190; Conservative 115; Mismatches 160; Indels 1; Gaps 1;

DB 61 IDXYIOLGYNNKNNVLYVGDVAVTNPVRKYKDGNEYIVKCKKKSQNNQADQVNA 120
QY 6 VNFILAMNDXKKLLTHGGESENEFKEGQLPXEYXXKRRKRSJSTNSDIXVAT 65
DB 121 ISSLTPLGALVYANSELVNPDVLPKRSLSLTIDPLGNTQDNKIVYKATRSYNN 180
QY 66 XDRSLRPLGLLVDETXTLNNPTLLAVDRPMTYSXPLGLASDSPFLQVEDPSSVRG 125
DB 181 AVNTLVERNNKXYAQAQYPRVSAKIDYDDDEMAYSQOLAKFETAKAYNNSLNVFGAIS 240
QY 126 AXDXDLAKMHODYGOVNVNPARXQYEXKTAHSEQLKVFSGPSEKXNSLIDFNSVH 184
DB 241 EGRMOEVIKFOIYVNVNNEPTRPSEFGKAVTKEOQLAGVAENPPAYISSVAGR 300
QY 185 SGKXIQVNNKQIYITVSDAVKNGVDYIVEDLQKRGISARPLVYISXVAYXR 244
DB 301 OYTKLSTNSHSTKYKAFDAVSKSVSGDVELTNIIRKNSFKAVIYGSAGKDEVQIID 360
QY 245 OYVLEKLETTSSXSEVEAFALIKGVKAPQTEWKQILNDNTYKAAVILGGDPSSGARVY 304
DB 361 GNLDLRLDKKATFNRTPGVPFAYTNTFLKNDLAVIKNSSEYITTKATTDGIN 420
QY 305 GKVADWEDLIOGSRFTADHPGLPISTYTFSLRDNVAVATFONSDYVEIKVATVRNGDL 364
DB 421 IDHSGEYVAGNYSWDEVNYPDEGNEIYQHKNSSENNKSLAHTSSYIPLGNAKNIN 480
QY 365 LDHSGAVYAOYITVWELSLDHOKEVLTFRKMRNODDLTAHFTYISPLKGVNRNL 424
DB 481 AKRECTGLAMEMWRTVYEDRNLPYKRNRISTIGTTLTPKYSNKNYDN 526
QY 425 IRECTGLAMEMWRTVYEKTDLXVKKRTISIMGTTLTPYVEDKVEN 470

RESULT 4
ID TACY LISIV STANDARD. PRT: 528 AA.
AC P31831;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE IVANOLYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).

```



```

QY 250 LETTSKSEVEAFALIKGVKAPQTEMKOIJDNTYKAVIIGLGDSSGCAVVTGKVM 309
DB 343 IQSVIKDNAQFSSKNRPAPISYTSVFLKDNSIAVHNNTETKTEYSKGIKIDHSG 402
QY 310 VEDLIQESRRTADHPGLPISYTSFLRDVNVATFQNSTDYVELTKVAYARNGLDLHSG 369
DB 403 AYVAQFEYVDEFSYDADGOEIVTRKSMWDRSARHFEETPLPPNKNKNIIFARECT 462
QY 370 AYVAQYITWELSYDHQKEVLTPKAMDRNGDLDIAHFTSLPLGNVRNLSVKIRECT 429
DB 463 GLAEMWRTVYDEVPALADINVTGTLTP 495
QY 430 GLAEMWRTVYKTDLXVLRKRTISWGTLTP 462

RESULT 6
ID TACY_LISSE STANDARD: PRT: 530 AA.
AC P31830:
DT 01-FEB-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE SEELIGERIOLYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN ISO.
OS LISTERIA SEELIGERI.
CC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
CC LISTERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SLCC:
RA MEDLINE: 92182018.
RX HAAS A., DUMBSKY M., KREFT J.;
RT "Listeriolysin genes: complete sequence of ilo from Listeria ivanovi1
and of ilo from Listeria seeligeri."
RL BIOCHIM. BIOPHYS. ACTA 1130:81-84(1992).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, X60462; G44145; -.
CC PIR, S22340; S22340.
CC DR PROSITE, PS00481; THIOL_CYTOLYSINS; 1.
CC DR PFAM, PF01289; ThiOL_cytolysin; 1.
CC DR HSSP, P19995; lPPO.
CC KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
CC FT SIGNAL 1 25
CC FT CHAIN 26 530 SEELIGERIOLYSIN.
CC FT ACT_SITE 485 485 BINDING TO CHOLESTEROL (BY SIMILARITY).
CC SEQUENCE 530 AA; 59181 MW; 68A497A5 CRC32;

Query Match 45.2%; Score 1414; DB 1; Length 530;
Best Local Similarity 40.6%; Pred. No. 1.17e-276;
Matches 189; Conservative 109; Mismatches 167; Indels 1; Gaps 1;
DB 62 INKYIWMGLNYSKNSLTVYOGSAVINPPKKYKDGSEIYVEKKKGGINONNADISVINA 121
QY 6 VNDFLIANNYDXXKLLTHQGESIENREKXENQOLPXEFEVXXEKRRKSLSTNTSDIYVXAT 65
DB 122 ISSLTIPGALVANKRELVEQNPVLPVKRDLSTLVLPMTKKDKNIIFKNFTKSNVNN 181
QY 66 XDSRLYPCGALLVDETXYENNPFTLAVDRAPMTYSXLPGLASSDSFLQVEDPSNSVSG 125
DB 182 AVNTLVEMWNRKYSKAIPIVNAKIDYSDENAYSSQILIAFGTAFAVNNLSLVNNEEAS 241

```

```

QY 126 AXDXLLAKMHDYGOVN-NVPAKQYKXKTAHSMEOIAKVFSGDFEKXGNSLDIDFNSVH 184
DB 242 DGAQOEVEISFKQIYYNNINNEPTSPKFFGGSYKREQLDALGVAENPVAISSVAYGR 301
QY 185 SGEXKIOIVAKKIOITYTVSDAVKNGDVEFQTVVEDLKQKIGISAEPRPLVAYSXVAYR 244
DB 302 QYVYKSSSSSHKVKVTAFAEAMSGKSVAGDELNTIINKNSFEKFIYVIGSGAKEVEIID 361
QY 245 QYVLKLETTSSXSXEVAFAEALIKGVKAPQTEMKOIJDNTYKAVIIGLGDSSGARVYT 304
DB 362 GNIGELRDLIAKSTSTDRNPGVPISTYTNELKNDLAVVKNNSIEYETSKSYTDGKIN 421
QY 305 GKVDWVEDLIQESRRTADHPGLPISYTSFLRDVNVATFQNSTDYVELTKVAYARNGLDL 364
DB 422 IDHSGYVAAQFNISMDVEYDENGNDIKVKKKMGKNYSKLAHFTSSLYLGNARNINITY 481
QY 365 LDHSGAVYVQYITWELSYDHQKEVLTPKAMDRNGDLDIAHFTSLPLGNVRNLSVK 424
DB 482 ARCTGLFEMWRTVYDEVPALADINVTGTLTPRHSNNVDN 527
QY 425 IRECTGLAEMWRTVYKTDLXVLRKRTISWGTLTPQVEADKVEN 470

RESULT 7
ID TACY_STRPY STANDARD: PRT: 571 AA.
AC P21131:
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS PYOGENES.
CC PLASMID PKM157.
CC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
CC STREPTOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE: 88057628.
RX KEROE M.A., MILLER L., WALKER J.A., BOUINOIS G.J.;
RT "Nucleotide sequence of the streptolysin O (SLO) gene: structural
RT homologies between SLO and other membrane-damaging, thiol-activated
RT toxins."
RL INFECT. IMMUN. 55:3228-3232(1987).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, M16638; G15811; -.
CC PIR, A43507; A43507.
CC DR PROSITE, PS00481; THIOL_CYTOLYSINS; 1.
CC DR PFAM, PF01289; ThiOL_cytolysin; 1.
CC DR HSSP, P19995; lPPO.
CC KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMID.
CC FT SIGNAL 1 33
CC FT CHAIN 34 571 STREPTOLYSIN O.
CC FT ACT_SITE 530 530 BINDING TO CHOLESTEROL (BY SIMILARITY).
CC SEQUENCE 571 AA; 63638 MW; 33124E54 CRC32;

Query Match 44.1%; Score 1379; DB 1; Length 571;
Best Local Similarity 40.7%; Pred. No. 8.80e-269;
Matches 186; Conservative 100; Mismatches 170; Indels 1; Gaps 1;
DB 109 INDKIYSLNNELEVLAKNGETIENFVPEKGVKADKFIYERKKKNINTTPVDISIIDS 168

```

Qy	6	VNDFILAMNYDXXKLLTHQGSIEINREKXEGKNOLPEFYXKERRKSUSTVTSDIYXAT	65
Dd	169	VYDTRPALOLANKGFENKPDAAVYTKRNPQKIHLDIGMDDKAT-VEVNDPYRANVT	227
Qy	66	XDSRLYPCGALLVDETXLENNPPTLADRAKPRMYKXKXLLGLASSOSFQYVEDPENNSSVRG	125
Dd	228	AIDNLVNMWHDNYSGGNTLPARTQYTESWYKSKQJLEALNWNKSLIDGTGICIDFKSTK	287
Qy	126	AXXDLIAKHQDQYGVGVNVPARXQYEXKTKASHMEDOLKVFSGDFPKXGNSJDIIDFNYSVHS	185
Dd	288	GEKKVIAIARQOIFETYSANLPNNPADVPFKSVYTFEKLOKRGKSNAPRLPYSNVAVGRT	347
Qy	186	GEKKIOIVNKKOIIYTVSDVANKNGDFEDQTVYEDLQRGISARPLPYVIXVAIKRQ	245
Dd	348	VFKLFTSKSDNVDFAEASALKGTIDVYNGKYSJLLENSFTAVLUGDAAEHNRKVVTK	407
Qy	246	VYTKLETTKSKXSEVERAEALIKGYKAPQTEMKQILDNTRYKAVILGSDPESSGARVVTG	309
Dd	408	DPDVIRNVIKDANATSRKNPAYPISTYSVFLKNKLAGVNRRTEYVETTSIEYISGKJNL	467
Qy	306	KVMDMEDLIOESRFLADHPGLPISTYSFLRDNVAVTFQNSDYEVETKYVYRANGDLL	365
Dd	468	SHOGAVYAQTEILIMDEINYDQKGVYTKRRKMDNNWYKSKSPSYIYPIGASRNRIRMA	522
Qy	366	DHSQAVYAQYIITWXLSDYHOGKREVLPRKAMDNRGODTLAHPFTTSIPLKGNVRNLSVKI	425
Dd	528	RECTGLAEMWKRVIDERPVYTKSEIINNYSIGRSLP	564
Qy	426	RECTGLAEMWKRIVYEKTDLKYRKRTISINGTLLP	462

RESULT ID	8	STANDARD:	PRF:	574 AA.
CC	TACY_SPREQ			
CC	054114:			
CC	15-JUL-1998 (REL. 36, CREATED)			
CC	15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)			
CC	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
CC	STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).			
CC	SLO.			
CC	STREPTOCOCCUS EGISIMILIS.			
CC	BACTERIA; FIRINICTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;			
CC	STREPTOCOCCUS.			
CC	[1]			
CC	SEQUENCE FROM N.A.			
CC	STRAIN-SIMD-1:			
CC	MEDLINE: 95102113.			
CC	OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MIAMIDE W.,			
CC	IGARASHI H., YUTSUDO T.,			
CC	"Cloning and sequencing the streptolysin O genes of group C and group			
CC	G streptococci.";			
CC	DNA SEQ. 4:325-328(1994).			
CC	-!- FUNCTION: SUEHYDRA-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL			
CC	CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.			
CC	CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO			
CC	EUKARYOTIC CELL MEMBRANES.			
CC	-!- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as this statement is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sb.ch/announce/			
CC	or send an email to license@isb-sb.ch).			
CC	-----			
CC	EMBL: D16824; G498301; -			
CC	PROSITE: PS00481; THIOL_CYTOLYSINS; 1.			
CC	PRAM: PF01289; Thiol_cytolysin; 1.			
CC	HSP: P19995; 1PFO.			
CC	TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.			
CC	SIGNAL			
CC	1			
CC	36			
CC	BY SIMILARITY.			
CC	STREPTOLYSIN O.			
CC	37			
CC	574			

FT	ACT_SITE	533	BINDING TO CHOLESTEROL (BY SIMILARITY)
SQ	SEQUENCE	574 AA; 63991 MM; 9BBF31B4 CRC32;	

44.18; Score 1380; DB 1; Length 574;

Query Match	44.1%	Score 1380;	DB 1;	Length 574;
Best Local Similarity	40.58;	Pred. No. 5,24e-269;		
Matches 185;	Conservative 102;	Mismatches 169;	Indels 1;	Gaps 1;

113 INDTYGI NYNET EVI ARNGET TENEVPRFEGVKKADKETVTERKKKNINTTPVDISIIDS 171

[illegible]

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

[illegible]

66 XDSRLYPGALLVDEIXLENNPILLAVDRAPMISXALFGLEHSDSE DQVEDEFNNOVNO

231 AIDNLVNQWHDNYSGGNTLPARTQYTESMYSKSQIEALNVNSKILDDGLGIDFKSLK 2

126 AXXDLAKWHQDYGVNNVPARXQYEKXTAHSMQLKVKFGSDFEKKXGNSLDIDENSVHS 1

291 GEKKVIAAYKQIFYTVSANLPNNPADVFEDKSVTLKELQRRKGVSNAPPLFVSNVAYGRT 3

```

186 CEKXOTLVNKKOTYYTTSVDANKPNPDVFODTVTVEDLKORGISAERPLVYISXVAYXRQ 2

```

086 NEWARK NJCENNYKNCVSDTI ENSSETAVVI GGDAEFHNKVTK 4

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

411 DEFVLRNKDNATFSKRNPAIPISISVFENNNNLRGVNNNCELVEISCELEDOONRANT
D6      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

306 KVDMEEDIQEGSRFTADHPGLPISYITSFLRDNVAALFQNSIDIVEIKVIAIRNGDLEL

Db 471 SHQGYVAQYEILWDEINYYDDKGKEVITKRWDNNWYSKTSFSTVIPLEGANSRNIRIMA 3

366 D HSGAYVAQYYITWXELSYDHQKEVLTPKAWDRNGQDLTAHFTTSIPLKGNVRLSVKI 4

Db 531 RECTGLAWEMWRKVIDERDVKLSKEINVNISGTLSP 567

00 436 PECTGI AWEFWRTVY EKTDLXI VRKRTISIWGTTLP 462

RESULT 9

ID	INCIDENT NO	DATE	TIME	LOCATION	DESCRIPTION	STATUS	ASSIGNED TO	RESOLUTION
AC	Q53957;							

DT	15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT	13-JUL-1998 (REL. 30, CANCELED)

DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 CTEPDTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN):

GN SLO.

BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;

```

OC SIKELVCCCCC:
RN [1]

```

RP	SEQUENCE FROM N.A.
RX	MEDLINE: 95102113.

OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
TAKAGUCHI Y., YOSHINO T.

RT "Cloning and sequencing the streptolysin O genes of group C and group

RL DNA SEQ. 4:325-328(1994).

-1- FUNCTION: SULFHYDRYL-ACTIVATED IOXIN. IS ABLE TO BIND AND CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION

CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE IONAINS TO
CC ERYTHROCYTIC CELL MEMBRANES

CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a colla

the European Bioinformatics Institute. There are no restrictions

CC use by non-profit institutions as long as its content is not modified and this statement is not removed. Usage by and for commercial purposes is prohibited.

CC entities requires a license agreement (See <http://www.isb-sid.cn/a>

2000 年 12 月 31 日 12 时 59 分 59 秒

```

CC -1- FUNCTION: CATALYZES THE OXIDATIVE DEAMINATION OF BIOGENIC AND
CC XENOBIOTIC AMINES AND HAS IMPORTANT FUNCTIONS IN THE METABOLISM OF
CC NEUROACTIVE AND VASOACTIVE AMINES IN THE CENTRAL NERVOUS SYSTEM
CC AND PERIPHERAL TISSUES. MAO-A PREFERENTIALLY OXIDIZES BIOGENIC
CC AMINES SUCH AS 5-HYDROXYTRYPTAMINE (5-HT), NOREPINEPHRINE AND
CC EPINEPHRINE.
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
CC H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: MONOMER, HOMO- OR HETERODIMER (CONTAINING TWO SUBUNITS OF
CC SIMILAR SIZE). EACH SUBUNIT CONTAINS A COVALENTLY BOUND FLAVIN.
CC ENZYMATICALLY ACTIVE AS MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL OUTER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X15609; G524; -.
CC DR EMBL: X15609; G525; ALT_INIT.
CC DR PIR: S03974; S03974.
CC KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; TRANSMEMBRANE; MITOCHONDRION;
CC NEUROTRANSMITTER DEGRADATION.
CC NP_BIND 14 70 FAD (ADP PART) (POTENTIAL).
CC FT BINDING 406 406 FAD (BY SIMILARITY).
CC FT TRANSMEM 498 518 POTENTIAL.
CC SQ SEQUENCE 527 AA; 59800 MW; 724B436 CRC32;

Query Match 3.6%; Score 112; DB 1; Length 527;
Best Local Similarity 18.7%; Pred. No. 6,96e-02;
Matches 28; Conservative 42; Mismatches 74; Indels 6; Gaps 5;

Db 1 MESLOKTSADGQMFVYVIGGSGISGISAAR-LIAEHVAVLVLEARERGATYVRNH 59
QY 250 LETTSXSEVEAEALVIGVAPQTEWKQILLDNVYKAVILGGPSSGAVYVTKKMDM 309
Db 60 V-YVYVVGAGVYPTQNRILRLSKGLFETKYVNERLVHVKGKTYFPRGAPPPVMP 118
QY 310 VEDLIQGSFTADHPCGLPISTYTSF-LRDNVATFQNSTDIYEVKTVYRNGDLLDHS 368
Db 119 IAYL-DYNNILRTM--DNNKKEIPADAPWE 145
QY 369 GAVYAOYIITWXLXSDHOGKEVLTPKAMD 398

RESULT 12
ID MPPB_YEAST STANDARD; PRT; 462 AA.
AC P10507;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE MITOCHONDRIAL PROCESSING PEPTIDASE BETA SUBUNIT PRECURSOR
DE (EC 3.4.24.64) (BETA-MP) (PEP).
GN MASI OR MIF1 OR YDL163C OR L9632.10.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88312592.
RA WITTE C., JENSEN R.E., YAFFE M.P., SCHATZ G.;
RT "Mas1, a gene essential for yeast mitochondrial assembly, encodes a
RT subunit of the mitochondrial processing protease.",
RL EMBL J. 7:1439-1447(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-S288C / AB972;
RC JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,

```

```

RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J.,
RA KOCABA T., HALLSWORTH K., HAWKINS J., HILLER L., JIER M.,
RA JOHNSTON D., JOHNSTON L., LANGSTON Y., LATREILLE P., LE T.,
RA NARDIS E., MENZES S., MILLER N., NHAN M., PAULEY A., PELUSO D.,
RA RIEKEN L., RILES L., TAICH A., TREVASIS E., VIGNATI D.,
RA WILCOX L., WOLDMAN P., VAUDIN M., WILSON R., WATERSTON R.;
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE OF 21-32.
RX MEDLINE; 91177897.
RA YANG M., GELI V., OPLIGER W., SUDA K., JAMES P., SCHATZ G.;
RT "The Mas-encoded processing protease of yeast mitochondria.
RT Interaction of the purified enzyme with signal peptides and a
RT purified precursor protein.",
RL J. BIOL. CHEM. 266:6416-6423(1991).
CC -1- FUNCTION: THE MITOCHONDRIAL PROCESSING PROTEASE CLEAVES
CC PRESEQUENCES FROM MITOCHONDRIAL PROTEIN PRECURSORS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF N-TERMINAL TRANSIT PEPTIDES FROM
CC PRECURSOR PROTEINS IMPORTED INTO THE MITOCHONDRIUM, TYPICALLY WITH
CC ARG IN POSITION P2.
CC -1- COFACTOR: REQUIRES DIVALENT CATIONS FOR ACTIVITY.
CC -1- SUBUNIT: HETERODIMER OF ALPHA AND BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
CC INSULINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X07649; G3887; -.
CC DR EMBL: U051921; G1234852; -.
CC DR PIR: S00552; S00552.
CC DR PIR: A38734; A38734.
CC SGD: L0001026; MASI.
CC DR PROSITE: PS00143; INSULINASE; 1.
CC KW PAM; PF00675; Insulinase; 1.
CC KM HYDROLASE; METALLOPROTEASE; ZINC; MITOCHONDRIUM; TRANSIT PEPTIDE.
CC FT TRANSIT 1 20 MITOCHONDRIAL PROCESSING PROTEASE
CC FT CHAIN 21 462 BETA SUBUNIT
CC FT METAL 70 70 ZINC (BY SIMILARITY).
CC FT ACT_SITE 73 73 ZINC (BY SIMILARITY).
CC FT METAL 74 74 ZINC (BY SIMILARITY).
CC FT METAL 150 150 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 462 AA; 51083 MW; 438D1777 CRC32;

Query Match 3.5%; Score 108; DB 1; Length 462;
Best Local Similarity 21.9%; Pred. No. 2.34e-01;
Matches 33; Conservative 45; Mismatches 63; Indels 10; Gaps 9;

Db 65 NNGTAHLEHLAEK-GTQ-NRSOGGLELEIENI-GSHLVATSRRENTVYAKSLQDIPK 121
QY 151 EKXTRAHMEQKAKYFGSDFKXGNSLIDFNVSHGEEKXIQIYNKQIYITYSV-DAYAN 209
Db 122 AVDLISDLTKSVLDNSAIRERDVIRRESEVDKMYDEVDHLHEITTKDOPLGRITL 181
QY 210 PGDFQDTIVIEDLKQNGISAERPIYVI-SX-VA--YXRQVYKL-ETTSXSEVEAAFE 264
Db 182 GPIKNKSTRIDLDKYITTKNYKGRNVLG 212
QY 265 ALIKGVKAVQTEWKQ-ILLDNVYKAVILGG 294

RESULT 13
ID SLAP_BACST STANDARD; PRT; 1228 AA.
AC P35825;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)

```

[illegible]

 W E S E L I
 (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_dp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 30 16:11:16 1999; Maspar time 28.94 Seconds

Tabular output not generated. 888,294 Million cell updates/sec

Title: >US-09-120-044-4
 Description: (1-471) from US09120044.pep
 Perfect Score: 3127
 Sequence: 1 MANKAVNDFILANNYDXKL.....TISWGTLLYPQVEDKVENND 471

Scoring table: PAM 150
 Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: sptrembl9

1:sp:archaea 2:sp:bacteria 3:sp:fungi 4:sp:human
 5:sp:invertebrate 6:sp:mammal 7:sp:mhc 8:sp:organelle
 9:sp:phage 10:sp:plant 11:sp:rodent 12:sp:unclassified
 13:sp:vertebrate 14:sp:virus

Statistics: Mean 49.293; Variance 92.130; scale 0.535

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1731	55.4	497	2	055996 SULLYSIN.	0.00e+00
2	1726	55.2	497	2	085102 HEMOLYSIN.	0.00e+00
3	1243	39.8	534	2	031241 PYOLYSIN.	2.24e-238
4	125	4.0	50	2	048772 LISTERIOLYSIN O (FRAGM	1.01e-03
5	119	3.8	50	2	048773 LISTERIOLYSIN O (FRAGM	7.42e-03
6	113	3.6	265	3	012044 CHROMOSOME XY READING	5.12e-02
7	113	3.6	265	3	068518 HIPOTHETICAL 49.7 KD P	5.12e-02
8	112	3.6	744	2	067108 DNA GYRASE A SUBUNIT.	7.01e-02
9	114	3.6	1361	2	087083 S-LAYER PROTEIN.	3.73e-02
10	114	3.6	1361	2	030524 S-LAYER PROTEIN.	3.73e-02
11	109	3.5	218	1	088691 GLUCOSE-6-PHOSPHATE IS	1.79e-01
12	110	3.5	550	2	085906 TRANSPOSON PROTEIN B.	1.31e-01
13	109	3.5	710	2	085906 TRANSPOSON PROTEIN B.	1.31e-01
14	108	3.5	814	10	065715 HIPOTHETICAL 91.9 KD P	2.43e-01
15	106	3.4	426	2	067578 FLAGELLAR HOOK ASSOCIA	1.10e+00
16	103	3.3	446	2	034363 PUTATIVE L-AMINO ACID	2.64e+00
17	100	3.2	142	10	049900 MEMIA (FRAGMENT).	3.52e+00
18	99	3.2	187	1	058072 187AA LONG HIPOTHETICA	3.52e+00
19	99	3.2	231	11	05813 KERATIN 14 (FRAGMENT).	2.64e+00
20	100	3.2	287	1	059566 287AA LONG HIPOTHETICA	2.64e+00

21	99	3.2	346	13	093404 PROLACTIN RECEPTOR (FR	3.52e+00
22	100	3.2	349	1	058654 HIPOTHETICAL PROTEIN M	2.64e+00
23	100	3.2	362	2	050434 AMINOTRANSFERASE.	2.64e+00
24	99	3.2	366	1	029257 ABC TRANSPORTER, ATP-B	3.52e+00
25	101	3.2	453	3	078803 ELUSION YEAST.	1.97e+00
26	101	3.2	501	1	028947 CONSERVED HIPOTHETICAL	1.97e+00
27	99	3.2	580	2	033559 TRANSDUCER-LIKE PROTEIN	3.52e+00
28	99	3.2	821	5	093378 C44H4.7 PROTEIN.	3.52e+00
29	99	3.2	877	3	066593 SIMILAR TO S. CEREVISIA	3.52e+00
30	99	3.2	903	3	087320 DYNAMIN-RELATED PROTEI	3.52e+00
31	99	3.2	1035	2	025887 CATION EFFLUX SYSTEM P	3.52e+00
32	99	3.2	1329	5	076356 C45G7.6 PROTEIN.	3.52e+00
33	100	3.2	1782	5	094682 MYOSIN V (DILUTE CLASS	2.64e+00
34	100	3.2	3027	11	063460 PROLINE-RICH PROTEIN (2.64e+00
35	96	3.1	281	3	092240 MULTICOPY SUPPRESSOR OF	8.24e+00
36	96	3.1	283	5	015923 DEOXYURIDINE TRIPHOSPH	6.22e+00
37	97	3.1	291	1	058761 HIPOTHETICAL PROTEIN M	6.22e+00
38	97	3.1	436	2	067031 TRANSCRIPTIONAL TERMIN	4.69e+00
39	96	3.1	437	2	074770 7-KETO-8-AMINOPELAGON	8.24e+00
40	96	3.1	498	5	017746 SIMILARITY TO C2H2-TYP	8.24e+00
41	96	3.1	498	5	068164 SENSOR PROTEIN BVRS.	8.24e+00
42	98	3.1	1532	13	090994 190 KD TENASCIN PRECUR	4.69e+00
43	98	3.1	1714	13	090995 200 KD TENASCIN PRECUR	4.69e+00
44	98	3.1	1810	13	090824 CYTOTACTIN 200KD PRECU	4.69e+00
45	97	3.1	2160	5	017709 H05L14.2 PROTEIN.	6.22e+00

ALIGNMENTS

RESULT ID	1	PRELIMINARY:	PRT:	497 AA.
AC	055996: 055997;			
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)			
DT	01-JAN-1999 (TREMBLREL. 09, LAST SEQUENCE UPDATE)			
DT	01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)			
DE	SULLYSIN.			
GN	SLY.			
OS	STREPTOCOCCUS SUTS.			
OC	BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;			
OC	STREPTOCOCCUS.			
RN	[1]			
RP	SEQUENCE OF 1-71 FROM N.A.			
RC	STRAIN-PL7/;			
RA	SEGERS R.P.A.M., KENTER T., DE HAAN W.A.M., JACOBS A.A.C.;			
RT	"Characterization of the gene encoding sullysin from Streptococcus			
RT	suts, and expression in field strains."			
RL	SUBMITTED (OCT-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; Z36907; E1334304; -			
SO	SEQUENCE 497 AA; 54850 MW; D78ABEF CRC32;			

Query Match 55.4%; Score 1731; DB 2; Length 497;
 Best Local Similarity 48.5%; Pred. No. 0.00e+00;
 Matches 227; Conservative 105; Mismatches 134; Indels 2; Gaps 2;

DB	29	SKQDINQFQSGITTPQELINNEGIDINPATIGMLENGFVVLRRKKNITNNSADIA	88
QY	2	AKKAVNDFILANNYDXKLTHQGESINREPKKNQLPXFEVXXEKRRKSLSTSDIX	61
DB	89	VLDAAKAIYFGALIRADONLNDNPTLISARGDLTSLNPLGLANGDSHTVNSPTRS	148
QY	62	VYAKIDSRLYIGALLVDEITLKNPITLAVDRAPMTYXKLPGLASSDSFLOVEDPSNS	121
DB	149	TYRTIVNNLLSKMNTTAGETGNTQALQYDETMAYSMSQLTFTGTSFKIAVPLDINF	208
QY	122	SVRGAXXDILAKWHDQY-GQVNNVNPARYQYKXTHAHSMEQKVFSGDFEXXGNSLIDF	180
DB	209	DAVNGEKOVOIVNPKOYIVSVDEBSPSKLAEGTYVEDLKRNGITDEPPVYSSV	268
QY	181	NSVHGEKXIQIVNKKQITIVSVDAVKNPDEVDQDVTYEDLKORISAEPLVYISXY	240
DB	269	SYGRSMFIKETSSRSITQVQAFAKIKGVDISGNAEYODILKNTSAYIFGGDGSAA	328
QY	241	AYXRGVYIKLETTXSKEVEAEFALLIKGVKVAPOETMKQILDNTYKAVIILGDPSGA	300

Query	Subject	Score	Length	Ident
329	TVVSGNIEITLKKIIEBARYGKLNPGPISTYFVGNQNPAAIISXEIETTSYVHNS	48.5%	497	329
301	RVVIGKAVMEDLLQIESRFPADHPGLPISITSLFDNVVATFONSDEIVETKVTAYRN	44.5%	495	301
389	SALTLDHSGAVVAYVNTTWEEVSNTEAGEEWEKPRANDKNGVNLTSMSSETIOPGNARN	44.5%	495	389
361	GDLLDHSAGVAAQYITVWXLSDHQGEVLTPKRAMDRNGODTLAHTFTSIPLKGNVRN	44.5%	495	361
449	LHVNIQECTGLAMEWRTIVDK-DLPVYGOKRITINGTITLYPOYADVEY	44.5%	495	449
421	LSVRIRECTGLAMEWRTIVETKDLXVKKRTISIMGTTITLYPOVEDKY	44.5%	495	421
085102	PRELIMINARY: PRT: 497 AA.			
085102	085102: PRELIMINARY: PRT: 497 AA.			
01-NOV-1998	(TREMBLREL. 08, CREATED)			
01-NOV-1998	(TREMBLREL. 08, LAST SEQUENCE UPDATE)			
01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)			
HEMOLYSIN				
STREPTOCOCCUS SUIS				
BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;				
STREPTOCOCCUS.				
(1)				
SEQUENCE FROM N.A.				
RC STRAIN-1933:				
RA OKUMABUA O.:				
"Streptococcus suis type 2 hemolysin (suisysin) gene";				
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.				
DR EMBL: AF043556; G3309561; "				
SQ SEQUENCE 497 AA; 54778 MW; 36DAB184 CRC32;				
Query Match	55.2%	Score 1726;	DB 2;	Length 497;
Best Local Similarity	48.5%	Pred. No. 0.00e+00;		
Matches 227;	Conservative 104;	Mismatches 135;	Indels 2;	Gaps 2
29	SKODINQYFOSLTLYGPOELITNEGEYIDNPATTGMLNGRFVYLRRKKNTNNSADIA	88		
2	ANKAVNDFILANNDYXXKLITHQGESIENRFKKEGNQCPXEFVXXERKRRLSTSTDX	61		
89	VIDAKAANTYPCALLRADONLDDNNPTLISAKGDLTSLNPLGLANDSHTVVNSPTRS	148		
62	VYATXDSRLPYALLVYDETLLNNPTLLAVDRAPMYTSXKLPLGLASSDSFLQVEDPSNS	121		
149	TVRTGVNNLLSKMNNVTAGEYGNFOALELOYDETMAYMSQSLTKRGTSEFKIAPLDINF	208		
122	SVRGAXXDLIAKMHODY-GQVNNVYARAKQYEXKXTHSHSQDLKYKRGSPFEKXGNSLDIDF	180		
209	DAVNSGEKQOYIVNFQKQIYTVSDPESSKFLAEGTTVEDLKRNGITIDEVPPYVSSV	268		
181	NSVHSEKKIQLVNNKKQIYTVSDVAVKNPGBVFDQTVTVEDLKORGISAERPLVYISXV	240		
269	SYGRSMFILTLETSRSTVOQAAPKAAIKQVDISGAAEYODILAKMSPFAYIFGDSGASAA	328		
241	AYXROYLLLETTSXSXEAAFEALIKQVAKAPOPTEMQKQIDLNTXKAVKALIGDPSGSA	300		
329	TVVSGNIEITLKKIIEBARYGKLNPGPISTYFVGNQNPAAIISXEIETTSYVHNS	388		
301	RVVIGKAVMEDLLQIESRFPADHPGLPISITSLFDNVVATFONSDEIVETKVTAYRN	360		
389	SALTLDHSGAVVAYVNTTWEEVSNTEAGEEWEKPRANDKNGVNLTSMSSETIOPGNARN	448		
361	GDLLDHSAGVAAQYITVWXLSDHQGEVLTPKRAMDRNGODTLAHTFTSIPLKGNVRN	420		
449	LHVNIQECTGLAMEWRTIVDK-DLPVYGOKRITINGTITLYPOYADVEY	495		
421	LSVRIRECTGLAMEWRTIVETKDLXVKKRTISIMGTTITLYPOVEDKY	468		
031241	PRELIMINARY: PRT: 534 AA.			

Query	Match	Score	DB	Length
Best Local	Similarity 38.1%;	Pred. No. 2,24e-238;		
Matches 180;	Conservative 118;	Mismatches 166;	Indels 8;	Gaps 6
Db	63	ENGUCUKYIRGLKYPDSGLAVKGSIEENVPTK-DOLKDGITYYFKHREKSFNNLRDIS	121	
Qy	3	NNAVDFILAMVYDXXKLTHQGESIENRFYKKNQDPXE-FVXKKRKSRLSTNTSDIX	61	
Db	122	AFDANNAHYPGALVLANKDILAKGSPISIGIARAPOIVSVDPGLVGDKNFVINVNPITS	181	
Qy	62	VKATXDSRLRYPCALLVDETXYLENNPPLLAVDRAPMYTSXXLPGLASDSFLOVEDPENS	121	
Db	182	SYTQGLINGLIDGWIOQRNSKYPDHAKISYDEDMYTSKOLLEAKIGLGEKYSAKLVNVED	241	
Qy	122	SVRGAXXDLDAWHDQVQNVNPARXQYEXKTXASHMQLVYKESDPEKKGNSLIDFN	181	
Db	242	AIHKREROVAIASFQIYYTASVDPDTPSHSVYGFNPVTAODLKRGVNNKNPLGYISSVS	301	
Qy	182	SVHSGKXKQIYNXKQIYYTASVDANKDGVDFQDVTVEDLKGKISABRPLYSXYA	241	
Db	302	YGRQIFVKEITTTSTNDVOAASGLFPAKFGNLSTEFKAKTADILNKTRATVYVANGSAR	361	
Qy	242	YKROYVLKLETTYSXSXEVAAPFAELIKG--YKVAPO--TEMKQILDNTYKAVIYIGDPS	297	
Db	362	GGVEVATGNDILKTIKEESTYSTEKVPAVPSYAVNLFKQNDQLAASVSSGDYIEITATT	421	
Qy	298	SGARVYTGAVDVEDDLIOGSRFTDHDGELPISYITSELRQNVAVTFQNSTDYVETKYTA	357	
Db	422	YKSGELTFRHGGGVYAKFLKMDIEISYDPQGEKITPTKWSGNMAARLLGRETIOQLAN	481	
Qy	358	YRNGDILLHSAIYAQYITWXLSTSDHQGEVLETPKRAMDNGODLTAHFETTSPLKGN	417	
Db	482	ARHIIHAEAGALGAMPWV-TVINKKNPLVPHREIYVKGTLNPNVEDNV	532	
Qy	418	VRLNVLSKIRECTGLAWE-WMTVYEKTDLXLPKRKRISIMGTILLPQYEDKV	468	
RESULT	4	PRELIMINARY:	PRT:	50 AA.
ID	048772			
AC	048772:			
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)			
DT	01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	LISTERIOLYSIN O (FRAGMENT).			
GN	HLV.			
OS	LISTERIA MONOCYTOGENES.			
OC	BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;			
OC	LISTERIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-11984, type 1;			
RC	MEDLINE: 96118685.			
RA	RASMUSSEN O.F., SKOUBOE P., DONS L., ROSSEN L., OLSEN J.E.;			

RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE: 98196666.
 RA DECKERT G., WARREN P.V., GASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBEER R., SNEAD M.A., KELLER M., AUCAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus";
 RL NATURE 392:353-358(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RA DECKERT G., WARREN P.V., GASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBEER R., SNEAD M.A., KELLER M., AUCAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
 RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AE000716; G2983485; -
 SQ SEQUENCE 744 AA; 84080 MW; 93795328 CRC32;

Query Match 3.6%; Score 112; DB 2; Length 744;
 Best Local Similarity 24.2%; Pred. No. 7.01e-02;
 Matches 22; Conservative 26; Mismatches 37; Indels 6; Gaps 5;

DB 449 IEY-YKKIVASGERT-KVFIEFEELVKKY--GD-KRTETIGVKKEKESITVAVIQD 503
 OY 13 MNYDXXKLLTHOGESIENRFKKGQNPXEFVXXERRKRSLSYNTSDIXXATYDSRLYP 72
 DB 504 GSIIPEELPLEKAP-VVNIILRVPTGGLFL 533
 OY 73 GALLVDETLENNPTLLAVDRAPMTYSXL 103

RESULT 9
 ID 087083 PRELIMINARY; PRT; 1361 AA.
 AC 087083;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE S-LAYER PROTEIN.
 DE CAMPYLOBACTER RECTUS.
 OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;
 OC CAMPYLOBACTER.
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33238;
 RX MEDLINE: 98442421.
 RA MIYAMOTO M., MAEDA H., KITANAKA M., KOKEGUCHI S., TAKASHIBA S.,
 RA MURAYAMA Y.;
 RT "The S-layer protein from Campylobacter rectus; sequence
 RT determination and function of the recombinant protein";
 RT J. BACTERIOL. 187:1563-1567(1998).
 DR EMBL: AB001876; D1034498; -
 KW S-LAYER.
 SQ SEQUENCE 1361 AA; 144904 MW; 55A9D736 CRC32;

Query Match 3.6%; Score 114; DB 2; Length 1361;
 Best Local Similarity 19.9%; Pred. No. 3.73e-02;
 Matches 37; Conservative 50; Mismatches 89; Indels 10; Gaps 9;

DB 178 VDMHSPHGRKAYELTTTNDATANVFENAPMK-HNPGGTDRIIMTLOSSDKLTGDSRHD 236
 OY 115 VEDPSNSVYRGAXXDLAKWHDYGOVNNVPAKXQYKXTHASME-QLKVFSGDFEXKG 173
 DB 237 NTLNVEFGANDEGDPISRTPLTINIONINIEVTGVNTLDL-RDSNDVEKINIRHTK 295
 OY 174 NSDIDIFNSVHS--GEKXIQIVNXXKQI--YYTVSDAVANPBGDFVFDVTDLXKRGISA 230
 DB 296 EAGNKNVNESIG-QKLYGMBLANVAK-KDIDVKEFH-KKGVLSGEDSKSNVLEAVEKS 352
 OY 231 ERPLVY-ISKVAXXROYVTKLETTSSKSEVEAFALIKGVAVAPOTEMKQILDITVYKA 289
 DB 353 LSITSD 358
 OY : : : : :
 CC : : : : :

OY 290 VILGSD 295

RESULT 10
 ID 030524 PRELIMINARY; PRT; 1361 AA.
 AC 030524;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE S-LAYER PROTEIN.
 DE CRS.
 GN CAMPYLOBACTER RECTUS.
 OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;
 OC CAMPYLOBACTER.
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-314;
 RX MEDLINE: 98187925.
 RA WANG B., KRAIG E., KOLODRUBETZ D.;
 RT "A new member of the S-layer protein family: characterization of the
 RT crs gene from Campylobacter rectus";
 RL INFECT. IMMUN. 66:1521-1526(1998).
 DR EMBL: AF010143; G2459961; -
 SQ SEQUENCE 1361 AA; 144385 MW; CAPE081F CRC32;

Query Match 3.6%; Score 114; DB 2; Length 1361;
 Best Local Similarity 19.9%; Pred. No. 3.73e-02;
 Matches 37; Conservative 50; Mismatches 89; Indels 10; Gaps 9;

DB 178 VDMHSPHGRKAYELTTTNDATANVFENAPMK-HNPGGTDRIIMTLOSSDKLTGDSRHD 236
 OY 115 VEDPSNSVYRGAXXDLAKWHDYGOVNNVPAKXQYKXTHASME-QLKVFSGDFEXKG 173
 DB 237 NTLNVEFGANDEGDPISRTPLTINIONINIEVTGVNTLDL-RDSNDVEKINIRHTK 295
 OY 174 NSDIDIFNSVHS--GEKXIQIVNXXKQI--YYTVSDAVANPBGDFVFDVTDLXKRGISA 230
 DB 296 EAGNKNVNESIG-QKLYGMBLANVAK-KDIDVKEFH-KKGVLSGEDSKSNVLEAVEKS 352
 OY 231 ERPLVY-ISKVAXXROYVTKLETTSSKSEVEAFALIKGVAVAPOTEMKQILDITVYKA 289
 DB 353 LSITSD 358
 OY 290 VILGSD 295
 OY : : : : :
 CC : : : : :

RESULT 11
 ID 058691 PRELIMINARY; PRT; 218 AA.
 AC 058691;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL PROTEIN M01295.
 GN M01295.
 OS METHANOCOCCUS JANNASCHII.
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
 OC METHANOCOCCUS.
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96337999.
 RA BOLT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
 RA STETTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
 RA KERRIDGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
 RA OVERBEER R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,
 RA SCOTT J.L., GEOHAGAN N.S.M., WEIDMAN J.F., FUHRMAN J.L., NGUYEN D.,
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKY M.,
 RA KLEIN H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii";
 RL SCIENCE 273:1058-1073(1996).
 CC -1- SIMILARITY: STRONG, TO A.FULGIDUS AF1550 AND
 CC M.THERMOAUTOTROPHICUM MTH576.

DR EMBL: U67570; G1591933; -
 KW HYPOTHETICAL PROTEIN.
 SC SEQUENCE 218 AA; 24666 MW; CFC99AB9 CRC32;
 Query Match
 Best Local Similarity 3.5%; Score 109; DB 1; Length 218;
 Matches 19; Conservative 12; Mismatches 29; Indels 2; Gaps 2;
 Db 86 KRAKIGAEIIVVH-GEIVVEPEEKTNYASISSEDVILAFPGFIDKETAENKENDIF 144
 QY 171 KXGNSLDIDFNSVHSGEKXIQVXKQIYTVSDAVKNPGDV-FQDTVYEDLKQRGIS 229
 Db 145 VE 146
 QY 230 AE 231
 RESULT 12
 ID 088015 PRELIMINARY; PRT; 550 AA.
 AC 088015.
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE GLUCOSE-6-PHOSPHATE ISOMERASE.
 GN PGI.
 OS STREPTOMYCES COELICOLOR.
 OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
 CC ACTINOMYCETALES; STREPTOMYCINAE; STREPTOMYCETACEAE; STREPTOMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA SEGER K.J., HARRIS D.;
 RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
 RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA MEDLINE; 97000351.
 RA REBENACH M., KIESER H.M., DENAPATITE D., EICHNER A., CULLUM J.,
 RA KINASHI H., HOPWOOD D.A.;
 RT "A set of ordered contigs and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL MOL. MICROBIOL. 21:77-96(1996).
 DR EMBL: AL031107; E1311951; -
 KW ISOMERASE.
 SO SEQUENCE 550 AA; 60424 MW; 028FD1F6 CRC32;
 Query Match
 Best Local Similarity 3.5%; Score 110; DB 2; Length 550;
 Matches 49; Conservative 53; Mismatches 104; Indels 17; Gaps 15;
 Db 245 ALSTNAEKVADFCDITNMEFMDWVGGRYSFSAIGLSLMIAGDPRFEMLDGFRIVD 304
 QY 52 SLSTNSIDIVXAXXSRLPGALLVVDXTLNNPTL-LAVDRAPMTYSXLPGL-AGS 109
 Db 305 EHRFNAEPANAPL--LLGLGWGDFLGAOSHVLPRSHLSLFTAY-LOQLMESN 360
 QY 110 DSFLQVEDPNSSSVRGAXXDLAKWHODY-G-OVNNVPAKXQY-EKXTAHSMQLKVF- 165
 Db 361 GKSYDREGNPNVOMQTGFWVG-TP-GT-NGQNAVYQLIHOGTKLIPADFLGARFVDEL 417
 QY 166 GSPFEKXGSLDIDFNSVHSGEKXIQVXKQIYTVSDAVK-NGGDVFODTIYVEDLK 224
 Db 418 E-ELKAOHDLMANFPA-QTQA-LAFGKTPDEVRAAGVPEELV 457
 QY 225 QRGISAERPLVYISVAYXKQVTLKLETTYSXSXEVAEFAFI 267
 RESULT 13

ID 085906 PRELIMINARY; PRT; 710 AA.
 AC 085906;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE TRANSPOSON PROTEIN B.
 GN TNPB.
 OS SPHINGOMONAS AROMATICIVORANS.
 OG PLASMID PNL1.
 OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; ZYMONOMAS GROUP;
 CC SPHINGOMONAS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-F199;
 RA ROMINE M.F., STILLMELT L.C., WONG K.-K., THURSTON S.J., SISK E.C.,
 RA SENSEN C.W., GASTERLAND T., SAFER J.D., FREDRICKSON J.K.;
 RT "Complete sequence of a 184 kb cataplastic plasmid from SpHINGomonas
 aromaticivorans strain F199."
 RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF079317; G3378338; -
 KW PLASMID.
 SO SEQUENCE 710 AA; 81079 MW; 41B48670 CRC32;
 Query Match
 Best Local Similarity 3.5%; Score 109; DB 2; Length 710;
 Matches 14; Conservative 20; Mismatches 25; Indels 3; Gaps 3;
 Db 139 NHHNDVLECYAAWYGANGR-ROSOT-ROMIRGKITYSEFSPVOLGDIRFAIT 196
 QY 366 DHGAVVAQYIYTWMLSTDHGKEVLTIPKANDRNGQDILAFHTTS-IPLKGWVNLISVK 424
 Db 197 NR 198
 QY 425 IR 426
 RESULT 14
 ID 065715 PRELIMINARY; PRT; 814 AA.
 AC 065715;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 91.9 KD PROTEIN.
 GN TSK18.220.
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 OC EUKARYOTA; VIRIDIPHYTES; EMBRYOPHYTA; TRACHEOPHYTA;
 OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
 CC CAPRALES; BRASSICACEAE; ARABIDOPSIS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BEVAN M., VAN DER SCHUEREN J., CHUANG Y.-J., VOET M., ROBBEN J.,
 RA VOLCKAERT G., BANCROFT I., MEWES H.W., MAYER K.F.X., SCHUELLER C.;
 RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU ARABIDOPSIS SEQUENCING PROJECT;
 RT SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AL022580; E1287635; -
 KW HYPOTHETICAL PROTEIN.
 SO SEQUENCE 814 AA; 91943 MW; A41B8987 CRC32;
 Query Match
 Best Local Similarity 3.5%; Score 108; DB 10; Length 814;
 Matches 21; Conservative 26; Mismatches 32; Indels 5; Gaps 5;
 Db 729 EMHSKNVHPKITYTVAIGYARGVNTFASRLNEMREKGIYPD-SITKKEFIYGLKQ 787
 QY 187 EKXIQVXKQIYTVSDAVKNPGDVFODTIYVEDLKQGISHERPLVYISV-AYXRO 245
 Db 788 GGV-LEAFKSGDE-EN-YAALIBG 808
 QY 246 VYKLETTYSXSXEVAEFAFI 269

```

RESULT 15
ID 067578 PRELIMINARY; PRT; 426 AA.
AC 067578;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE FLAGELLAR HOOK ASSOCIATED PROTEIN FLK.
GN FLK.
OS AQUIFEX AECOLICUS.
OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RN [1]
RC STRAIN-VF5;
RX MEDLINE: 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aecolicus."
RL NATURE 392:353-358(1998).
RN [2]
RC SEQUENCE FROM N.A.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AE000749; G2983977; -.
KW FLAGELLA.
SQ SEQUENCE 426 AA; 49567 MW; 47417AE0 CRC32;

Query Match 3.4%; Score 106; DB 2; Length 426;
Best Local Similarity 23.7%; Pred. No. 4.47e-01;
Matches 22; Conservative 24; Mismatches 39; Indels 8; Gaps 7;

Db 1 MEGASENTISOLEV-YKKM-IDVKNRNINAAQENTVAEEPVQSDLYSGITFQEVRI 58
OY 164 KFGSPDEKXGNSLDIDFNSVHSGEKXIOIVXKQIYYTVSYDAVK-N-PGDVFQDTVTY 220
Db 59 QNFNFINTRNK-LSYVYLE-ERRDYLSKLES 89
OY 221 EDLKRGISAEPLVYISXVAYKROYL-KLET 252

```

Search completed: Mon Aug 30 16:12:37 1999
 Job time : 81 secs.


```

Db 181 nsvhsgkqigvlnfkqiyvtsvdknpgdvfgdltvvelkqrgisaezplyissv 240
|
|
|
Oy 181 NSVHSGEKQIOIYNKKQIYVTSVDAVKNPBGVFOPTVVEDLKRGISAEERPLYISXV 240
241 AYGRQVYLKLETTSKSXEVFAEALIKGVKAPQTEWKQIIDNTEVKAVILGDPSSGA 300
|
|
|
Oy 241 AGRQVYLLKLETTSKSXEVFAEALIKGVKAPQTEWKQIIDNTEVKAVILGDPSSGA 300
301 RVVTGKVDWEDLIQSGSFTADHPGLPISTTSLRDVAVTFQNSDYEVTATARN 360
|
|
|
Db 301 RVVTGKVDWEDLIQSGSFTADHPGLPISTTSLRDVAVTFQNSDYEVTATARN 360
361 gdlldhsagayvaqyyitwnelsydhqkevltpkamdngodlrahfTTSIPKGNRN 420
|
|
|
Oy 361 GDLLDHSGAYVAQYYITWNELSYDHQKEVLTPKAMDNGODLRAHFHTTSPKGNRN 420
421 LSVKIRECTGLAMWMTVYEKTDLPVKKRTISWGTTLPQVEDKEND 471
|
|
|
Oy 421 LSVKIRECTGLAMWMTVYEKTDLPVKKRTISWGTTLPQVEDKEND 471

```

RESULT 2

```

ID R05924 standard; protein; 471 AA.
AC R05924;
DE 23-NOV-1990 (first entry)
DE Immunogenic pneumolysin variant.
KW Pneumolysin; vaccine; pneumonia; meningitis; bacteraemia; ds.
OS Streptococcus pneumoniae.
FH Key Location/Qualifiers
FT misc_difference 428
FT FT /label=g, A, S
FT FT misc_difference 435
FT FT /label=Q, D
PN WO9006951-A.
PE 28-JUN-1990.
PD 15-DEC-1989; AU0539.
PR 16-DEC-1988; AU-001989.
PA (PATO/ PATON J C.
PI PATON JC, HANSMAN DJ, BOULNOIS GJ, ANDREW PW, MITCHELL TJ,
PI WALKER JA;
PI WPI; 90-224494/29.
DR New non-toxic, immunogenic mutants of pneumolysin - useful in
PT protective vaccines against Streptococcus pneumoniae, and DNA
PS sequences encoding them.
PS Disclosures: P: English.
CC Vaccines are non-toxic and antigenic to wild type pneumolysin,
CC making them useful in vaccination against pneumonia and associated
CC meningitis, bacteraemia etc.
SQ Sequence 471 AA;

```

Query Match 94.5%; Score 3063; DB 1; Length 471;
 Best Local Similarity 94.7%; Pred. No. 2, 07e-257;
 Matches 446; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

```

Db 1 mankvndflamoydkkkllthgesientfikegnlqdeefviekksistnsdi 60
|
|
|
Oy 1 MANKVNDFILAMNYDKKKLLTHGESIENTFRKXKGNQLPDEFVYXERKKKSLSTNSDI 60
61 svlatndsrlypsaallvvdellennptllavdrapmtysldlpglaasdsflgvedpsn 120
|
|
|
Oy 61 XVTATNDSRLYLGALLVVDDELLENPTLLAVDRAPMTYSIDLPGLASDSDFLOVEDPSN 120
121 ssrvtaavndllakwbqdygqvnnparmqyekkhtahsmeglkvfkfsgsfektgnsldidf 180
|
|
|
Oy 121 SSVRGAHVNDLLAKWHQDYGOVNNVPARQYKHTAHSMEOQLKVFEGSDFEKTSGLDIDF 180
181 nsvhsgkqigvlnfkqiyvtsvdknpgdvfgdltvvelkqrgisaezplyissv 240
|
|
|
Oy 181 NSVHSGEKQIOIYNKKQIYVTSVDAVKNPBGVFOPTVVEDLKRGISAEERPLYISXV 240
241 AYGRQVYLKLETTSKSXEVFAEALIKGVKAPQTEWKQIIDNTEVKAVILGDPSSGA 300
|
|
|

```

```

Oy 241 AGRQVYLLKLETTSKSXEVFAEALIKGVKAPQTEWKQIIDNTEVKAVILGDPSSGA 300
|
|
|
Db 301 RVVTGKVDWEDLIQSGSFTADHPGLPISTTSLRDVAVTFQNSDYEVTATARN 360
301 RVVTGKVDWEDLIQSGSFTADHPGLPISTTSLRDVAVTFQNSDYEVTATARN 360
|
|
|
Db 361 gdlldhsagayvaqyyitwnelsydhqkevltpkamdngodlrahfTTSIPKGNRN 420
|
|
|
Oy 361 GDLLDHSGAYVAQYYITWNELSYDHQKEVLTPKAMDNGODLRAHFHTTSPKGNRN 420
421 LSVKIRECTGLAMWMTVYEKTDLPVKKRTISWGTTLPQVEDKEND 471
|
|
|
Oy 421 LSVKIRECTGLAMWMTVYEKTDLPVKKRTISWGTTLPQVEDKEND 471

```

RESULT 3

```

ID R3842 standard; Protein; 480 AA.
AC R3842;
DE 15-JUL-1993 (first entry)
DE Streptolysin O variant mSLO.3/6.
KW SLO; soluble; haemolytic activity; wild type; anti-SLO antibodies;
KW Streptococcus pyogenes; ASO.
OS Streptococcus pyogenes.
PN WO9305155-A.
PD 18-MAR-1993.
PE 03-AUG-1992; U06380.
PR 30-AUG-1991; US-752428.
PA (BECL ) BECKMAN INSTR INC.
PI Adams CW;
PI WPI; 93-100979/12.
DR N-PSDB; Q38287.
PT Streptolysin O variants produced by recombinant DNA technology -
PT having no haemolytic activity and recognised by wild-type
PT anti-streptolysin O antibodies, useful in diagnosis of
PT Streptococcus pyogenes infection
PS Claim 17; Fig 2; 80bp; English.
CC The sequence represents a soluble variant of Streptolysin O (SLO)
CC having no haemolytic activity. The variant SLO can specifically
CC recognise wild-type anti-SLO antibodies (ASO) but has no haemolytic
CC activity. It can be used for the determination of previous and/or
CC current infection by Streptococcus pyogenes. It enables ASO assays
CC to become commercially viable and avoids the risk of handling haemo-
CC lytically active S. pyogenes wild-type SLO.
SQ Sequence 480 AA;

```

Query Match 43.9%; Score 1423; DB 6; Length 480;
 Best Local Similarity 42.2%; Pred. No. 1, 96e-111;
 Matches 193; Conservative 104; Mismatches 159; Indels 1; Gaps 1;

```

Db 18 indklyslnynlevalakngetienvfvpkgykkadkfivlerkkknintpvdslids 77
|
|
|
Oy 6 VNDFILAMNYDKKKLLTHGESIENTFRKXKGNQLPDEFVYXERKKKSLSTNSDI 65
78 vdrtypaalqglanbgfctenkpdavvtkrnpklnhdipgmqdkat-vevndplyanvt 136
|
|
|
Oy 66 NDSRLYLGALLVVDDELLENPTLLAVDRAPMTYSIDLPGLASDSDFLOVEDPSNSVYRG 125
137 aldnlyngndhysgntltpartytksmvysksglaealnvnskllgtdgdfktsk 196
|
|
|
Oy 126 AVNDLLAKWHQDYGOVNNVPARQYKHTAHSMEOQLKVFEGSDFEKTSGLDIDFNSVHS 185
197 gekkmiiaeykqiflytsanlpnpadvfdksvtfkelqkrkyvsnearplfvsnvaygrt 256
|
|
|
Oy 186 GEKQIOIYNKKQIYVTSVDAVKNPBGVFOPTVVEDLKRGISAEERPLYISXVAVGRQ 245
257 vfkletsksndveaafaalkydvtkngkysdlensstlavvjgdaehnkvvtk 316
|
|
|
Oy 246 VYKLETTSKSXEVFAEALIKGVKAPQTEWKQIIDNTEVKAVILGDPSSGARVVTG 305
317 dfaivrvkdnatsrknpayisvsvflknkilaqvnrtteyvetsteysgkln 376
|
|
|
Oy 306 KYDMWEDLIQSGSFTADHPGLPISTTSLRDVAVTFQNSDYEVTATARN 365

```


[illegible]

PT negative bacterium - or new sbs-B gene in any host, also new
PI recombinant proteins containing heterologous inserts, e.g.
PT epitope(s), useful as vaccines and adjuvants
PS Claim 1; Pages 9-14; 31pp; German.
CC The present sequence is the *Bacillus stearothermophilus* PV72
CC S-layer protein, sbs-A. S-layer structures can be used as vaccines
CC or adjuvants, particularly when they include a bacterial ghost that
CC may contain additional epitopes in its membrane. Other uses of
CC recombinant sbs-A, depending on the nature of the inserted peptide,
CC are as an universal carrier for bioconjugated reactants for use in
CC immunological or hybridisation assays (the insert is streptavidin),
CC to induce immune responses (epitopes), as a reagent for removing
CC cytokine or toxin from serum (antigenic epitopes), as a molecular
CC spinning nozzle (polyhydroxybutyrate synthase), as a molecular
CC laser (luciferase).
SQ Sequence 1228 AA;

RESULT 11
ID W55347 standard; Protein; 188 AA.
AC W55347;
DT 17-JUN-1998 (first entry)
DE H. pylori ORF 01cel1618orf11 protein.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS Helicobacter pylori.
PN MO9737044<AL.
PD 00

RESULT 11
ID W55347 standard; Protein: 188 AA.
AC W55347;
DT 17-JUN-1998 (first entry)
DE H. pylori ORF 01ccl1618orf1 protein.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
PN MO937044.1.
PD 09-OCT-1997.
PE 27-MAR-1997; U05223.
PR 06-DEC-1996; US-761318.
PR 29-MAR-1996; US-625811.
PR 02-APR-1996; US-758731.
PR 25-OCT-1996; US-736905.
PR 28-OCT-1996; US-738859.
PR (ASTR) ASTRA AB.
PI Alm RA, Smith D;
DR WPI; 97-503122/46.
DR N-PSDB; V24756.
PT Helicobacter pylori nucleic acid sequences and encoded
PT polypeptide(s) useful in vaccines to treat or prevent H. pylori
PT infection and for diagnosis of H. pylori infection
PS Claim 14; Page 571; 1145pp; English.
CC This sequence is a H. pylori protein of unspecified function.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors. The
CC DNA and probes derived from it may be used for the identification of
CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
CC acid sequences complementary to the DNA act as antisense sequences and
CC can be used to prevent the translation of H. pylori mRNA. Antibodies
CC against the protein can be used in immunoassays to evaluate the abundance
CC and distribution of H. pylori-specific antigens. The genomic sequence of
CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
CC by mechanically shearing the bacterial DNA. The sequences were analysed
CC for ORF of at least 180 nucleotides, and the predicted coding regions
CC defined by computer evaluation. To identify likely H. pylori antigens for

CC vaccine development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or exported
CC membrane proteins. Having identified and determined the sequences of
CC interest, particular regions can be isolated from *H. pylori* by PCR
CC amplification for recombinant polypeptide production, e.g. in *E. coli*
CC hosts.
SQ Sequence 188 AA;

Query Match 3.3%; Score 108; DB 29; Length 188;
Best Local Similarity 34.4%; Pred. No. 1.12e+01;
Matches 21; Conservative 20; Mismatches 17; Indels 3; Gaps 3;

Dd 95 psalmrdsvsledskkr-1naikdlfhnkaf-rq1qlklnp1kal-veaqkdgefka 151
Y 210 PGDFVFDITVEDLKRGISAEPLVYSXVAYGRQVYIKLETSKSEVEAFALING 269
Dd 152 1 152
Y 270 V 270

RESULT 12
ID W60976 standard; protein; 299 AA.
AC W60976; (first entry)
DT 13-OCT-1998
DE Streptococcus pneumoniae encoded polypeptide.
KW coding region; ORF; open reading frame; antibacterial;
KW infection; prevention; meningitis.
OS Streptococcus pneumoniae.
PN W09819689-A1.
PD 14-MAR-1998
PF 27-OCT-1997; U19226
PR 01-NOV-1996; US-029930.
PA (SMK) SMITHKLINE BEECHAM CORP.
PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO,
PI Reid RH, Zarfos PN;
PI WPI; 98-285586/25.
DR N-PSDB; V37377.
PT New isolated nucleic acids from Streptococcus pneumoniae - useful,
PT e.g. for identifying anti-bacterial(s) for treatment and prevention
PT of meningitis.
PS Claim 11; Page 71-72; 130pp; English.
CC The sequence is that of the polypeptide encoded by a region isolated
CC from *S. pneumoniae* which shows homology to nicotinate-nucleotide pyro-
CC phosphoriase. It, or agonists of it, may be useful as an anti-
CC bacterial for treatment or prevention of infection, specifically caused
CC by *S. pneumoniae* (particularly meningitis) but possibly also *Helicobacter*
CC pylori (ulcers and gastric cancer). It may be of particular
CC use before insertion of an in-dwelling device or any other
CC invasive procedure. The protein, or nucleic acid encoding
CC it, can also be used in vaccines to induce a cellular
CC and/or humoral immune response, or to screen for other
CC antibacterials. The DNA may also contain flanking sequences
CC that are potential sources of control elements for bacterial
CC gene expression. Detecting a sequence encoding the protein
CC can be used diagnostically, e.g. to detect a mutation for
CC serotyping or classifying infectious agents.
SQ Sequence 299 AA;

Query Match 3.3%; Score 108; DB 33; Length 289;
Best Local Similarity 21.4%; Pred. No. 1.12e+01;
Matches 22; Conservative 35; Mismatches 40; Indels 6; Gaps 6;
Dd 38 dyahtaidhggakvsifakeagvlag-1tvfgvftlidaevtfqnpqfdgdrils 96
Y 14 NDKKRLTHGCESTENFKEGNDLPDEFVXERKRSLSTMTS-DIXVATNDSRLYP 72
Dd 97 gdl-vl-eiig-svrslltcervalnfhqisgiasmaayea 136
Y 73 GALLVDETLKENPTLLAVADRAPMPTYSIDLPGAS-SDSFLQ 114

RESULT 13
ID W55528 standard; protein; 418 AA.

AC W55528; (first entry)
DT 02-JUL-1998
DE *H. pylori* ORF 29ge30321_24336712-fl-5 cellular protein.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS *Helicobacter pylori*.
PN W09737044-A1.
PD 09-OCT-1997.
PF 27-MAR-1997; U05223.

PR 06-DEC-1996; US-761318.
PR 29-MAR-1996; US-625811.
PR 02-APR-1996; US-758731.
PR 25-OCT-1996; US-736905.
PR 28-OCT-1996; US-738859.
PA (ASTR) ASTRA AB.
PI Alm RA, Smith D;
PI WPI; 97-503122/46.
DR N-PSDB; V24937.
PT *Helicobacter pylori* nucleic acid sequences and encoded
PT polypeptides) - useful in vaccines to treat or prevent *H. pylori*
PT infection and for diagnosis of *H. pylori* infection
PS Claims 14, 93; Page 730-731; 1145pp; English.

CC This sequence is a *H. pylori* cellular protein.
CC The protein may be used in a vaccine to prevent or treat *H. pylori*
CC infection or to identify *H. pylori* polypeptide binding compounds, The
CC infection or to identify *H. pylori* life cycle activators or inhibitors. The
CC useful as potential *H. pylori* life cycle activators or inhibitors. The
CC DNA and probes derived from it may be used for the identification of
CC *H. pylori* in a sample and the diagnosis of *H. pylori* infection. Nucleic
CC acid sequences complementary to the DNA act as antisense sequences and
CC can be used to prevent the translation of *H. pylori* mRNA. Antibodies
CC against the protein can be used in immunoassays to evaluate the abundance
CC and distribution of *H. pylori*-specific antigens. The genomic sequence of
CC *H. pylori* (ATCC 55679) was determined from overlapping contigs generated
CC by mechanically shearing the bacterial DNA. The sequences were analysed
CC for ORF of at least 180 nucleotides, and the predicted coding regions for
CC defined by computer evaluation. To identify likely *H. pylori* antigens for
CC vaccine development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or exported
CC membrane proteins. Having identified and determined the sequences of
CC interest, particular regions can be isolated from *H. pylori* by PCR
CC amplification for recombinant polypeptide production, e.g. in *E. coli*
CC hosts.
SQ Sequence 418 AA;

Query Match 3.3%; Score 108; DB 29; Length 418;
Best Local Similarity 34.4%; Pred. No. 1.12e+01;
Matches 21; Conservative 20; Mismatches 17; Indels 3; Gaps 3;

Dd 325 psalmrdsvsledskkr-1naikdlfhnkaf-rq1qlklnp1kal-veaqkdgefka 381
Y 210 PGDFVFDITVEDLKRGISAEPLVYSXVAYGRQVYIKLETSKSEVEAFALING 269
Dd 382 1 382
Y 270 V 270

RESULT 14
ID W61278 standard; protein; 527 AA.
AC W61278; (first entry)
DT 29-SEP-1998
DE Monamine oxidase A.
KW Monamine oxidase A; abnormal behaviour; serotonin; dopamine;
KW noradrenaline; mental retardation; impulsive aggression.
OS Homo sapiens.
PN US5783680-A.
PD 21-JUL-1998.
PF 06-OCT-1993; 132168.
PR 06-OCT-1993; US-132168.
PA (GENO) GEN HOSPITAL CORP.

PA (UYKA-) UNIV STICHTING KATHOLIEKE.
 PI Breakfield XO, Brunner HG;
 DR WPI: 98-427102/36.
 DR N-PSDB: V27798.
 PT Genetic diagnosis of impulsive aggression - uses mutant forms of
 PS monoamine oxidase enzyme
 PS Disclosure: Column 47-50; 30pp; English.
 CC A genetic defect linked to p11-p21 of the X chromosome, especially
 CC mutation C936T in exon 8 of the monoamine oxidase (MAO) gene is related to
 CC abnormal behaviour. MAO metabolises serotonin, dopamine and
 CC noradrenaline, neurotransmitters whose metabolism is known to affect
 CC behaviour. Disregulation of these transmitters leads to borderline
 CC mental retardation and abnormal behaviour (including impulsive
 CC aggression). Probes based on the mutated sequence can be used for the
 CC genetic diagnosis of abnormal behaviour.
 SQ Sequence 527 AA;

Query Match 3.1%; Score 99; DB 33; Length 527;
 Best Local Similarity 19.3%; Pred. No. 4.32e+01;

Matches 29; Conservative 40; Mismatches 75; Indels 6; Gaps 5;

DB 1 mengekasiagmfdvvgvlgsgisaak-llteygsvlvleardvgrtyltineh 59
 OY 250 LETTSKSEVEAFEPALIKGVKAPOTEMKQILDNTVEKAVILGDPSSGARVYTKYDM 309
 DB 60 v-dydvvgagayvptqnrlrlskelgietykvnseilvyvkgktyfgrafpvpwmp 118
 OY 310 VEDLIOEGSRFTADHPGLPISYTSFLRDNV-VATFQNSTDYETKVTAYRNGDLLDHS 368
 DB 119 layl-dynmlwrti--dmngkeiptdapwe 145
 OY 369 GAVVAQYITWNELSYDHQKEVLTPEKAMD 398

RESULT 15

ID R05079 standard; protein: 527 AA.

AC R05079:

DT 10-MAR-1993 (revised)

DT 11-JUL-1990 (first entry)

DE Human monoamine oxidase type A gene product.

KW Human monoamine oxidase type A; MOA-A; mental retardation;

KW manic depression; psychotic disorders.

OS Homo sapiens.

PN MO9000195-A.

PD 11-JAN-1990.

PF 29-JUN-1989; 02901.

PR 30-JUN-1988; US-213544.

PA (BREA) Breakfield X.

PI Breakfield X.

DR WPI: 90-037130/05.

DR N-PSDB: Q03217.

PT Purified DNA encoding human monoamine oxidase type A -

PT used to detect eg. MAO-A alterations, associated with manic

PT depression, and expressed to form MAO inhibitors to treat psychotic

PT disorders.

PS Disclosure: Fig 2; 30pp; English.

CC MAO-A catalyses oxidative deamination of dietary amines and

CC neurotransmitters, eg. dopamine. The product may be used to develop

CC therapeutic inhibitors, to metabolise dietary monoamines or to act as

CC false transmitters.

SQ Sequence 527 AA;

Query Match 3.1%; Score 99; DB 2; Length 527;

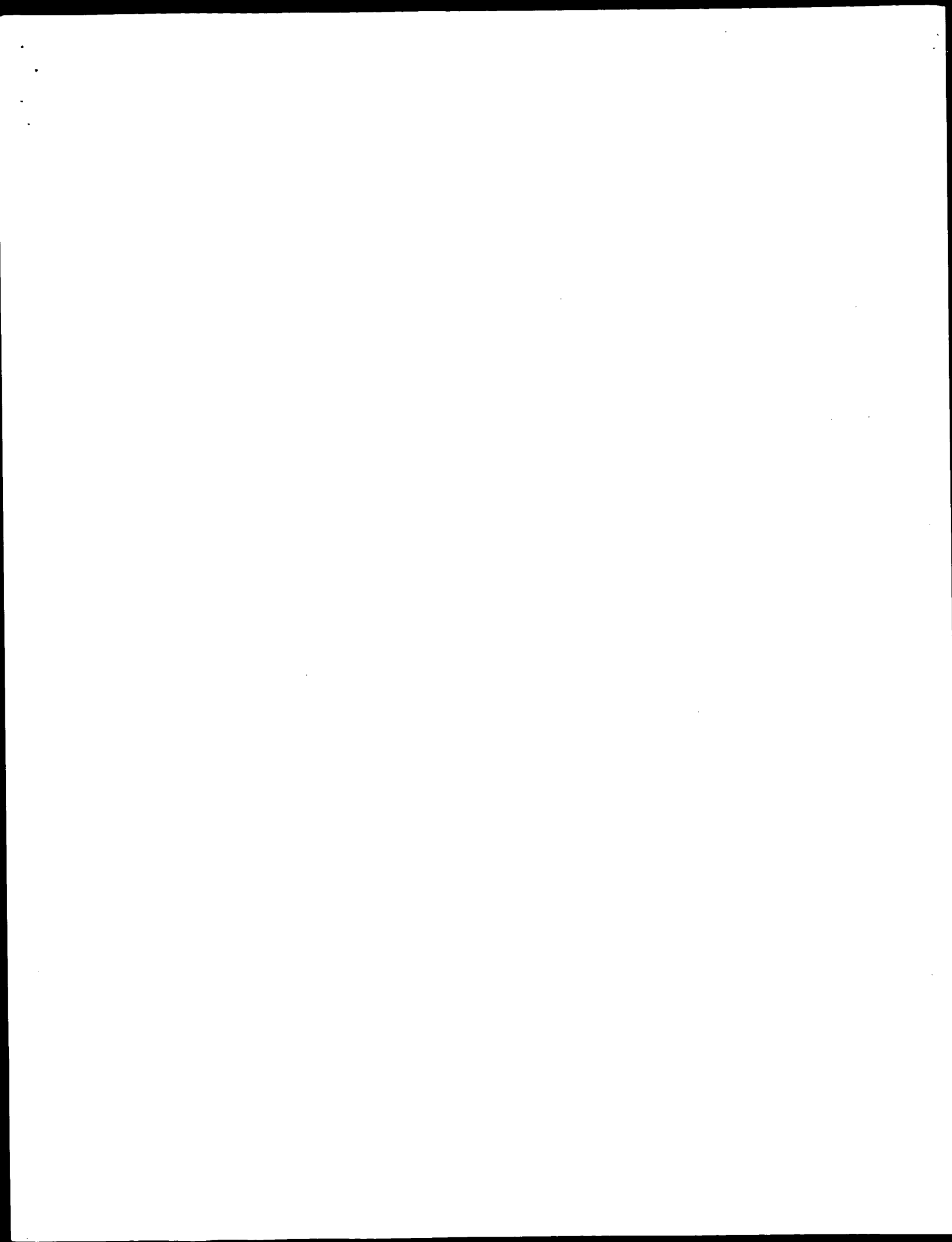
Best Local Similarity 19.3%; Pred. No. 4.32e+01;

Matches 29; Conservative 40; Mismatches 75; Indels 6; Gaps 5;

DB 1 mengekasiagmfdvvgvlgsgisaak-llteygsvlvleardvgrtyltineh 59
 OY 250 LETTSKSEVEAFEPALIKGVKAPOTEMKQILDNTVEKAVILGDPSSGARVYTKYDM 309
 DB 60 v-dydvvgagayvptqnrlrlskelgietykvnseilvyvkgktyfgrafpvpwmp 118
 OY 310 VEDLIOEGSRFTADHPGLPISYTSFLRDNV-VATFQNSTDYETKVTAYRNGDLLDHS 368

DB 119 layl-dynmlwrti--dmngkeiptdapwe 145
 OY 369 GAVVAQYITWNELSYDHQKEVLTPEKAMD 398

Search completed: Mon Aug 30 16:04:20 1999
 Job time : 114 secs.



Db 121 SSVRGAVNDLLAKKHODYGVNNVPAKQYKRTAHSMEOUKYKSGDFEKTGSLDIDF 180
 QY 121 SSVRGAVNDLLAKKHODYGVNNVPAKQYKRTAHSMEOUKYKSGDFEKTGSLDIDF 180
 Db 181 NSVHSGEKOIOIVNFKOQIYTVSDAVAKNPGDFQDVTYVEDLKORGISAEPLVYISV 240
 QY 181 NSVHSGEKOIOIVNFKOQIYTVSDAVAKNPGDFQDVTYVEDLKORGISAEPLVYISV 240
 Db 241 AYGRQVYVLTETSKSDVEAEAFALIKGVYAPQTEWKQILDNTYKAVYLGSDPSSGA 300
 QY 241 AYGRQVYVLTETSKSDVEAEAFALIKGVYAPQTEWKQILDNTYKAVYLGSDPSSGA 300
 Db 301 RYVTKVMDVEDLIOGSRFTADHPGLPISTTSLRNVAVATONSDYETKVTAYRN 360
 QY 301 RYVTKVMDVEDLIOGSRFTADHPGLPISTTSLRNVAVATONSDYETKVTAYRN 360
 Db 361 GDLILDHSGAVYAOYIITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPLKGVNRN 420
 QY 361 GDLILDHSGAVYAOYIITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPLKGVNRN 420
 Db 421 LSVKIRECTGLAMEWRYEKTDLPLVKKRTISITWGTTLTYQVBDKXEND 471
 QY 421 LSVKIRECTGLAMEWRYEKTDLPLVKKRTISITWGTTLTYQVBDKXEND 471
 RESULT 2
 ENTRY B43577 #type complete
 TITLE perfringolysin O precursor - Clostridium perfringens
 ORGANISM #formal_name Clostridium perfringens
 DATE 03-Feb-1993 #sequence_revision 03-Mar-1993 #text_change 13-Sep-1998
 ACCESSIONS B43577; A34951; A60922
 REFERENCES Shimizu, T.; Okabe, A.; Minami, J.; Hayashi, H.
 Infect. Immun. (1991) 59:137-142
 #journal An upstream regulatory sequence stimulates expression of the perfringolysin O gene of Clostridium perfringens.
 #title translation of the nucleotide sequence is not complete
 #cross-references MUID:91099951
 #accession B43577
 #molecule_type DNA
 #residues 1-500 #label SHI
 #cross-references GB:M61080; NID:g144885; PID:g144886
 #note Translation of the nucleotide sequence is not complete
 REFERENCE A34951
 #authors Tweten, R.K.
 #journal Infect. Immun. (1988) 56:3235-3240
 #title Nucleotide sequence of the gene for perfringolysin O (theta-toxin) from Clostridium perfringens: significant homology with the genes for streptolysin O and pneumolysin.
 #cross-references MUID:89032623
 #accession A34951
 #molecule_type DNA
 #residues 1-125, 'EA', 129-500 #label TWE
 #cross-references GB:M6704; NID:g144883; PID:g144884
 REFERENCE A60922
 #authors Tweten, R.K.
 #journal Infect. Immun. (1988) 56:3228-3234
 #title Cloning and expression in Escherichia coli of the perfringolysin O (theta-toxin) gene from Clostridium perfringens and characterization of the gene product.
 #cross-references MUID:85032622
 #accession A60922
 #molecule_type protein
 #residues 29-45 #label TW2
 #experimental_source ATCC 13124
 GENETICS pfo; pfca
 FEATURE 1-28 #domain signal sequence #status predicted #label SIG
 29-500 #product perfringolysin O #status experimental #label MAT
 SUMMARY #length 500 #molecular_weight 55799 #checksum 4821

Query Match 49.3%; Score 1599; DB 2; Length 500;
 Best Local Similarity 46.5%; Pred. No. 1,85e-270;
 Matches 214; Conservative 105; Mismatches 140; Indels 1; Gaps 1;

Db 35 NOSIDSGJSSLSYNNNEVLASNGDKIESFVPEKKGAKNRTIYERORSLTSPVISI 94
 QY 35 NOSIDSGJSSLSYNNNEVLASNGDKIESFVPEKKGAKNRTIYERORSLTSPVISI 94
 Db 95 IDSVNDRTYPGALADRAFAVENRPTIIMVKKRPININIDPLGKGENS-IKVDPTIYK 153
 QY 95 IDSVNDRTYPGALADRAFAVENRPTIIMVKKRPININIDPLGKGENS-IKVDPTIYK 153
 Db 63 TATNDRILYFALLVYDELAKENPTLLAVBPAPMTYSIDPLGLASSDSFQVDPSSNS 122
 QY 63 TATNDRILYFALLVYDELAKENPTLLAVBPAPMTYSIDPLGLASSDSFQVDPSSNS 122
 Db 154 VSGAIDELVSKNNEKISSSTHILPARTQYSESVMYKSSQISSALNNAVLENSIGVDENA 213
 QY 154 VSGAIDELVSKNNEKISSSTHILPARTQYSESVMYKSSQISSALNNAVLENSIGVDENA 213
 Db 123 VRGAVNDLLAKKHODYGVNNVPAKQYKRTAHSMEOUKYKSGDFEKTGSLDIDNS 182
 QY 123 VRGAVNDLLAKKHODYGVNNVPAKQYKRTAHSMEOUKYKSGDFEKTGSLDIDNS 182
 Db 214 VANNEKKMILAVKOEFYVSADLPKRNPSDLDPDSVFRTNDLKOKSVNEADPLMVSNAV 273
 QY 214 VANNEKKMILAVKOEFYVSADLPKRNPSDLDPDSVFRTNDLKOKSVNEADPLMVSNAV 273
 Db 183 VHSGEKOIOIVNFKOQIYTVSDAVAKNPGDFQDVTYVEDLKORGISAEPLVYISXVAY 242
 QY 183 VHSGEKOIOIVNFKOQIYTVSDAVAKNPGDFQDVTYVEDLKORGISAEPLVYISXVAY 242
 Db 274 GRTIYVLETTSSSKDVOAFKALIKNDIKNSQYKDIYENSFTAYVYLGDAQENHKV 333
 QY 274 GRTIYVLETTSSSKDVOAFKALIKNDIKNSQYKDIYENSFTAYVYLGDAQENHKV 333
 Db 243 GRQVYVLTETSKSDVEAEAFALIKGVYAPQTEWKQILDNTYKAVYLGSDPSSGARV 302
 QY 243 GRQVYVLTETSKSDVEAEAFALIKGVYAPQTEWKQILDNTYKAVYLGSDPSSGARV 302
 Db 334 VTKDPDEIRKRVYDKNATFESTKNPAPISYTSVFLKONSVAAHNKTVDYIETSTYSKAK 393
 QY 334 VTKDPDEIRKRVYDKNATFESTKNPAPISYTSVFLKONSVAAHNKTVDYIETSTYSKAK 393
 Db 303 VTKVMDVEDLIOGSRFTADHPGLPISTTSLRNVAVATONSDYETKVTAYRN 362
 QY 303 VTKVMDVEDLIOGSRFTADHPGLPISTTSLRNVAVATONSDYETKVTAYRN 362
 Db 394 INDHSGAVYAOYIITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPLKGVNRN 453
 QY 394 INDHSGAVYAOYIITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPLKGVNRN 453
 Db 363 LLDHSGAVYAOYIITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPLKGVNRN 422
 QY 363 LLDHSGAVYAOYIITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPLKGVNRN 422
 Db 454 IKARECTGLAMEWRYEKTDLPLVKKRTISITWGTTLTY 493
 QY 454 IKARECTGLAMEWRYEKTDLPLVKKRTISITWGTTLTY 493
 RESULT 3
 ENTRY A43505 #type complete
 TITLE listeriolysin O precursor - Listeria monocytogenes
 ORGANISM #formal_name Listeria monocytogenes
 DATE 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 17-Mar-1999
 ACCESSIONS A43505; S05306; A47606; S12400; A61079
 REFERENCES Mengaud, J.; Vicente, M.F.; Chenevert, J.; Pereira, J.M.; Georffroy, C.; Gicquel-Sauze, B.; Baquero, F.; Perez-Diaz, J.C.; Cossart, P.
 Infect. Immun. (1988) 56:766-772
 #journal Expression in Escherichia coli and sequence analysis of the listeriolysin O determinant of Listeria monocytogenes.
 #title listeriolysin O determinant of Listeria monocytogenes.
 #cross-references MUID:88153053
 #accession A43505
 #status preliminary
 #molecule_type DNA
 #residues 1-529 #label MEN
 #cross-references GB:M24199; NID:g149652; PID:g149653
 #note this sequence is derived from a strongly hemolytic strain, serotype 1/2c
 REFERENCE S05306
 #authors Domann, E.; Chakraborty, T.
 #journal Nucleic Acids Res. (1989) 17:6406
 #title Nucleotide sequence of the listeriolysin gene from a Listeria monocytogenes serotype 1/2a strain.
 #cross-references MUID:88366684
 #accession S05306
 #molecule_type DNA
 #residues 1-529 #label DOM
 #cross-references EMBL:X15127; NID:g44106; PID:g44107
 #experimental_source strain EGD
 #note this sequence is derived from a weakly hemolytic strain, serotype 1/2a
 REFERENCE A47606

```

#authors      Mengaud, J.; Chenevert, J.; Geoffroy, C.; Gallard, J.L.;
#journal      Infect. Immun. (1987) 55:3225-3227
#title        Identification of the structural gene encoding the
              SH-activated hemolysin of Listeria monocytogenes:
              listeriolysin O is homologous to streptolysin O and
              pneumolysin.
#cross-references MIMD:88057627
#accession    A47606
#status       preliminary
#molecule-type DNA
#residues     413-480 #label ME2
#cross-references GB:M29171
#accession    S12400
#authors      Michel, E.; Reich, K.A.; Favier, R.; Berche, P.; Cossart, P.
#journal      Mol. Microbiol. (1990) 4:2167-2178
#title        Attenuated mutants of the intracellular bacterium Listeria
              monocytogenes obtained by single amino acid substitutions
              in listeriolysin O.
#cross-references MIMD:91211627
#accession    S12400
#molecule-type DNA
#residues     483-493 #label MTC
#experimental_source strain LO28, serotype 1/2c

GENETICS
#gene          hlyA; hlyA
#classification #superfamily dipeptide transport protein
#keywords       virulence factor
#feature        1-25
#summary        #domain signal sequence #status predicted #label SIG\
                #product listeriolysin O #status predicted #label MAT
                #length 529 #molecular-weight 58688 #checksum 719

Query Match      47.6%; Score 1542; DB 2; Length 529;
Best Local Similarity 42.9%; Pred. No. 1,62e-259;
Matches 200; Conservative 117; Mismatches 148; Indels 1; Gaps 1;

Db 61 IDKTYGDLNKNVLYHGDAVTNPPRRGYKDGNEYIVKRRKKSINONNADIOYVNA 120
    :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 6 VNDFTLANNYDKKKLLTHGSEIENFRKXGNOQLPDEFVYXERKRSLSNTSDIXVAT 65
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
121 ISSLTYPGALVANKSELVENQDVLVPRKDSLTSLSDLPGMTNODKRIYVKNATKSNVNN 180
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
66 NDSRLYPGALLVYDETLKXENPFTLLAVDRAPMTYSIDLPGLAASSDFLOVEDPSNSVVG 125
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
181 AVNTLVERNNEKYAQAAPVNSAKIDYDEMAVSESOLAKFGTAFAVNNSLNVNGAIS 240
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
126 AVNDLLAKWHQDYGVN-NVPAKQYERKLTASHMEDLAKYFGSDPEKGTNSLIDFNSVH 184
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
241 EGMKEEVISFOQIYVNVNNEPTRPREFGKAVTKEOALGVNAENPPAYISSVAYGR 300
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
185 SGEKQIQIVNKKQIYTVSDAVKNPGDVFQDTIYEDLKQGISAEPLVYISXVAYGR 244
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
301 QVYLKLTNSHSTKKAAPDAVSGKSVGDELNIINNSFKAVIYIGSAKDEVOIID 360
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
245 QVYLKLTSTSKSXEVAEALIKGVKAPQTEMQOILNTEVKAVILGGDPSGCAVYT 304
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
361 GNLGDLRLIKGATFNRETPGVPIAYTTNKLKDNLAIVKNNSEIETTSKAYTDGKIN 420
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
305 GAVDMEDELIDGSRFTADHPGLPISTTSFRLDNVAFQNSDYVEKTVATARNQDL 364
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
421 IDHSGVYAOENISWDEINYPDEGNEIYOHKNWSENNKSKLAHFTSSYLPGNARNINVY 480
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
365 LDHSGAVYAOYITTNELSYDHQGREVLTPKAMDNGODLTAHFTTSLPLKGNVRLNSVK 424
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
481 AKECTGLAMEMWRTVIDDRNLPLVKNRNIISMGTTLYPKYSNKNYDN 526
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
425 IRECTGLAMEMWRTVYEKTDLPVYKRRTISIMGTTLYPOVEDKVEN 470
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 4
ENTRY   S24231 #type complete
TITLE   listeriolysin precursor - Listeria monocytogenes (strain

```

```

ORGANISM      12067)
#formal_name  Listeria monocytogenes
#strain       12067
#variety      22-Nov-1993 #sequence_revision 10-Nov-1995 #text-change
DATE          13-Sep-1998

ACCESSIONS
#authors      S24231
#journal      S24230
#title        Rasmussen, O.F.; Beck, T.; Olsen, J.E.; Dons, L.; Rossen, L.
              Infect. Immun. (1991) 59:3945-3951
              Listeria monocytogenes isolates can be classified into two
              major types according to the sequence of the listeriolysin
              gene.
#cross-references MIMD:92040062
#accession    S24231
#status       nucleic acid sequence not shown; translation not shown
#molecule-type DNA
#residues     1-529 #label RAS
#cross-references EMBL:X60035; NID:944110; PID:944112
#experimental_source strain 12067, serotype 4b
#note          the nucleotide sequence was submitted to the EMBL Data
              Library, June 1991

GENETICS
#gene          hlyA
#classification #superfamily dipeptide transport protein
#keywords       virulence factor
#feature        1-25
#summary        #domain signal sequence #status predicted #label SIG\
                #product listeriolysin #status predicted #label MAT
                #length 529 #molecular-weight 58687 #checksum 47

Query Match      47.4%; Score 1535; DB 2; Length 529;
Best Local Similarity 42.7%; Pred. No. 3.58e-258;
Matches 199; Conservative 117; Mismatches 149; Indels 1; Gaps 1;

Db 61 IDKTYGDLNKNVLYHGDAVTNPPRRGYKDGNEYIVKRRKKSINONNADIOYVNA 120
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 6 VNDFTLANNYDKKKLLTHGSEIENFRKXGNOQLPDEFVYXERKRSLSNTSDIXVAT 65
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
121 ISSLTYPGALVANKSELVENQDVLVPRKDSLTSLSDLPGMTNODKRIYVKNATKSNVNN 180
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
66 NDSRLYPGALLVYDETLKXENPFTLLAVDRAPMTYSIDLPGLAASSDFLOVEDPSNSVVG 125
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
181 AVNTLVERNNEKYAQAAPVNSAKIDYDEMAVSESOLAKFGTAFAVNNSLNVNGAIS 240
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
126 AVNDLLAKWHQDYGVN-NVPAKQYERKLTASHMEDLAKYFGSDPEKGTNSLIDFNSVH 184
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
241 EGMKEEVISFOQIYVNVNNEPTRPREFGKAVTKEOALGVNAENPPAYISSVAYGR 300
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
185 SGEKQIQIVNKKQIYTVSDAVKNPGDVFQDTIYEDLKQGISAEPLVYISXVAYGR 244
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
301 QVYLKLTNSHSTKKAAPDAVSGKSVGDELNIINNSFKAVIYIGSAKDEVOIID 360
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
245 QVYLKLTSTSKSXEVAEALIKGVKAPQTEMQOILNTEVKAVILGGDPSGCAVYT 304
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
361 GNLGDLRLIKGATFNRETPGVPIAYTTNKLKDNLAIVKNNSEIETTSKAYTDGKIN 420
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
305 GAVDMEDELIDGSRFTADHPGLPISTTSFRLDNVAFQNSDYVEKTVATARNQDL 364
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
421 IDHSGVYAOENISWDEINYPDEGNEIYOHKNWSENNKSKLAHFTSSYLPGNARNINVY 480
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
365 LDHSGAVYAOYITTNELSYDHQGREVLTPKAMDNGODLTAHFTTSLPLKGNVRLNSVK 424
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
481 AKECTGLAMEMWRTVIDDRNLPLVKNRNIISMGTTLYPKYSNKNYDN 526
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
425 IRECTGLAMEMWRTVYEKTDLPVYKRRTISIMGTTLYPOVEDKVEN 470
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 5
ENTRY   S22341 #type complete
TITLE   ivanolysin precursor - Listeria ivanovi
ORGANISM #formal_name Listeria ivanovi
DATE     07-Apr-1994 #sequence_revision 07-Apr-1994 #text-change
        21-Aug-1998

```

ACCESSIONS	S22341; S36683
REFERENCE	S22340
#authors	Haas, A.; Dumbsky, M.; Krefit, J.
#journal	Biochim. Biophys. Acta (1992) 1130:81-84
#title	Listeriolysin genes: complete sequence of 110 from Listeria Ivanovii and of 180 from Listeria seeligeri.
#cross-references	MUID:92182018
#accession	S22341
##molecule_type	DNA
##residues	1-528 ##label HAS
##cross-references	EMBL:X60461
#note	the authors translated the codon ACA for residue 331 as Val
REFERENCE	S36683
#authors	Krefit, J.
#submission	submitted to the EMBL Data Library, July 1991
#accession	S36683
##molecule_type	DNA
##residues	1-319, 'T', 321-528 ##label KRE
##cross-references	EMBL:X60461
GENETICS	
#gene	llo
CLASSIFICATION	#superfamily dipeptide transport protein
FEATURE	
1-24	#domain signal sequence #status predicted #label SIG\
25-528	#product lvanolysin #status predicted #label MAT
SUMMARY	#length 528 #molecular_weight 58511 #checksum 6874
Query Match	47.1% Score 1525; DB 2; Length 528;
Best Local Similarity	42.9%; Pred. No.2,976-286;
Matches	200; Conservative 118; Mismatches 147; Indels 1; Gaps 1
Db	60 IDQYIGDIDYDKNNILVYDEGAVKNVPKAGYKESGYIVYEKKKKSJNQNADIQVINS 119
Oy	6 VNDFTIAMYDKDKLLTHQGESIEENRFKEKGQLDEVEYERKKRSJSTSDIXATYAT 65
Db	120 LASLYAPGLAVKANSELVNOPDPVLPVKRDSVTLSIDLPGMWNHNEIVGNATKSNIND 179
Oy	66 NDSRLYPGALLVVDTEKLKNNPTTLAANDRAPRTYSITDLPGCLASDSPFLQVEDPSSSVRG 125
Db	180 GVNTLVDRNNNKSYSEEPNISAKIDYDDEMAYSESOLVAKFGAAEKAVNSLNWFGAIS 239
Oy	126 AVNDLLAKWHQDYGVN-NVPAKQYEREITASMELQAKKFESDFEKTGNSLDIDFNSVH 184
Db	240 ECKVOEEVINFOITYTVNVNNEPTSPSRFEKSYKENQALGVANENPPAYISVAAGR 299
Oy	185 SGKKOIOYXKQIOITYTVSDAVKANGDVFDQTVVEDLKORGISAERPPLYISKVALGR 244
Db	300 DIPVKLTSSHSSTRVAAADAFAFKGSVGDPELENIIONASGFKAIVYGSAKDEVEIID 359
Oy	245 QYYLKLETTSKSEVAEAERALIKGVKAPQREMQRIDNTEEVKAILGGDPSSGARVYT 304
Db	360 GDLSLRDLIKOGANKFNPKGPVPLVATTNFELKDNO LAVYKKNSEYIEFTSKAYSDDGIN 419
Oy	305 GKADWDIEDLOESRFTADHPGIPISYTFSLPDNVAAFQNSTDVERKYATVYNGDLL 364
Db	420 LDHSGVAVENFTWMEVSVDANGNVEVHKKSEMDOKLAHFTTSIILPGNARNINIH 479
Oy	365 LDHSGAVVAOYITTNWELSTDHGKREVLPRKAMDNRGODLTIAHTFSIDPKKNVNLSVK 424
Db	480 AKECTGLAWEMRTVVDDRNLPLVKNRNVCINGTLLTPAYSIDTVDN 525
Oy	425 IRECTGLAWEMRTVYEKTDLPVRKRRTISIWGTLLTPQVEDKVEN 470
RESULT	6
ENTRY	S22340 #type complete
TITLE	seeligeriolysin - listeria seeligeri
ORGANISM	#formal_name Listeria seeligeri
DATE	22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 21-Aug-1998
ACCESSIONS	S22340
REFERENCE	S22340

#authors	Haas, A.; Dumbsky, M.; Kieft, J.
#journal	Biochim. Biophys. Acta (1992) 1130:81-84
#title	Listeriolysin genes: complete sequence of ilo from <i>Listeria ivanovii</i> and of ilo from <i>Listeria seeligeri</i> .
#cross-references	NCBI:92182018
#accession	S22340
#status	preliminary
#molecule_type	DNA
#residues	1-530 #label HAA
#cross-references	EMBL:X60462; NID:q44144; PID:q44145
#note	the authors translated the codon GCC for residue 287 as Pro
CLASSIFICATION	#superfamily dipeptide transport protein
SUMMARY	#length 530 #molecular-weight 59181 #checksum 4523
Query Match	45.5%; Score 1476; DB 2; Length 530;
Best Local Similarity	42.5%; Pred. No. 7.35e-247;
Matches	198; Conservative 111; Mismatches 156; Indels 1; Gaps 1
Db	62 INKIMGVLDKNSILVYQGEAVTNPPKKGYKDGSEYIYVEKKKGINQONNADISVINA 121
QY	6 VNFILAMNDKKRLTLHQGESIENRFKKENQNPDEFEVYKKRKSLSLTNSDXVAT 65
Db	122 ISSLTLPGALVKANRELVEQNPVLPYKRDLSLTSLVDPGMTKKDKKIFVKNPTKSNVNN 181
QY	66 NDSRLTPGALLVYDELTKENNPGLLVANDRAMYISIDPLASDSDFLOVEDPSSNSVVG 125
Db	182 AVNTLVERMNDKSKAYPNINAKIDYSDEMAYSSQLIAFGTAFAVNVNSLVNFEALS 241
QY	126 AVNDLLAKHMDDQGVGV-NVPAKRYEKETITASHSEQLKVFQSDPEKFTNSDIDFNSVH 184
Db	242 DGKQOEVSISFRQIYVNVNPEPTSPKFFGSGVTKFCOLDALVNNENPPAYISSVAVGR 301
QY	185 SGEKQIDIVAKKQIYTVTSVDVAKKPGDVPDQIVTVEDLKQKRGISMERPLVIYSVAAGR 244
Db	302 QVYVKKSSSSHKVKTAFEAEMASGKSVKGVDELNTNIRKNSFRKAVIYGSAGEVEIID 361
QY	245 QVYLKLETTSSKSEVAEAFELIKGVKVAPOTEKQKOLDITVEKAVILGGDSSSGARVVT 304
Db	362 GNLGELDLILKKSITDRENPEVPISTYTNLKNQDLVAVKKNSEYIEFTSKSYTDGKIN 421
QY	305 GKVMVWEDLLQESRFTADPGPLSTYTSFLRDVAVATPQNSIDYETIKVYAIRNGDDL 364
Db	422 IDHGGVAAOFNISMPDEVSIDENGEMIEVHKWGENYKSKLAHFTSSILPGNARNINITY 481
QY	365 LDHSGAVAAQYIITMELSLSDHQGEKVLVPRKAMDRNGDGLAHFTSLPLKGNVRLSVK 424
Db	482 ARECTGFEMWRVIVDDRLPLVKNRNVSTMGTLVYRHSNNNDN 527
QY	425 IRECTGLAEMWRVIVEKTDLPVLRKRITISMGTLVPOVEDKVEN 470
RESULT	7
ENTRY	A37858 #type complete
TITLE	alveolysin - <i>Bacillus alvei</i>
ORGANISM	#formal_name <i>Bacillus alvei</i>
DATE	28-Jun-1991 #sequence_revision 28-Jun-1991 #text-change 09-Sep-1997
ACCESSIONS	A37858
REFERENCE	A37858
#journal	Geoffroy, C.; Mengaud, J.; Alouf, J.E.; Cossart, P.
#title	J. Bacteriol. (1990) 172:7301-7305
	Alveolysin, the thiol-activated toxin of <i>Bacillus alvei</i> , is homologous to listeriolysin O, pertingolysin O, pneumolysin, and streptolysin O and contains a single cysteine.
#cross-references	NCBI:91072224
#accession	A37858
#status	preliminary
#molecule_type	DNA
#residues	1-501 #label GEO
#cross-references	CB:M62709; NID:q142472; PID:q142473
SUMMARY	#length 501 #molecular-weight 55268 #checksum 8775

Query Match 45.3%; Score 1468; DB 2; Length 501;
 Best Local Similarity 41.9%; Pred. No. 2,51e-245;
 Matches 190; Conservative 121; Mismatches 141; Indels 1; Gaps 1;

DB 44 INGLNRRREVLAIGDDSSSEVPKKEGINSNGKFIYVERDKKSLTSPVDISIVDSITNR 103
 10 IIAANNDRKKLLTHOGESIEENRFKXEGNOLPDEFVYXERKKRSLSTNTSDIXVATNDSR 69
 DB 104 IYPGAIALANKDFADNOPFLVMAARKPLDISIDLPGLKNENT-IVQNPNGTSSAIDQ 162
 70 LKPGALLVDEFLXENNPFLAVDAPMTYSIDLPGLASSDSFLOVEDPSNSVGAAND 129
 DB 163 LVSTGEKYSSTHTLPARLOYAESMYSONQISSALYNNAKVLNGLIDENAVANGEEK 222
 130 LIAKHMHODYGVANNVPARKQYEKITAHSMEOLKVFSGDEFKTSGLDIDRNSHSGKQ 189
 DB 223 VWAAYAKOFEYTVSAGLPNNPSDLDDSYTFAELARKGVSNAPPLMYSNVAAGRTIYVK 282
 190 IOIVAKKQIYTVSVDAVKNPGDVFODTYVEDLKORGISAEPLVYISXVAYGROYLX 249
 DB 283 LETTSKNDVQTAFLKLLNPNPSIQASGOYKDIYENSSTFAVLGGDAQTHNOYTKEDNV 342
 250 LETTSKSEVEAEFALLKGVAVAPQTEMKQILDNTYKAVILGGDPSSGARVYTGKAYDM 309
 DB 343 IQSVTKDNAGSSKNPAPVISTYSVFLKDNSTIAAVHNTYETETKTEYSKGIKLDHSG 402
 310 VEDLQESRFTADHPGLPISYTSFLRDNVAVAFQNSTDYVERKVTYVRNGDLDLHDSG 369
 DB 403 AYVAOFEYVMEFSDADGOEIVTKPSWDGMRBSAFSTEIPLPNAKIRIFARECT 462
 370 AYVAOYITWMELESTDHQKEVLTPKAWDRNGODLTAFTTISIPLKGNVRLSVKIRECT 429
 DB 463 GLAEMWRTVYDEVYVPLASDINVSIMGTTLYP 495
 430 GLAEMWRTVYEKTDLPVYRKRTISIMGTTLYP 462

RESULT 8
 ENTRY A43507 #type complete
 TITLE streptolysin O precursor - Streptococcus pyogenes
 ORGANISM #formal name Streptococcus pyogenes
 DATE 28-Oct-1992 #sequence_revision 28-Oct-1992 #text_change 09-Sep-1997

ACCESSIONS A43507
 REFERENCE A43507
 #authors Kehoe, M.A.; Miller, L.; Walker, J.A.; Boulinois, G.J.
 #journal Infect. Immun. (1987) 55:3228-3232
 #title Nucleotide sequence of the streptolysin O (SLO) gene: structural homologues between SLO and other membrane-damaging, thiol-activated toxins.

#cross-references M01D:88057628
 #accession A43507
 #status preliminary
 #molecule_type DNA
 #residues 1-571 #label KEH
 #cross-references GB:M18638; NID:9153810; PID:9153811
 SUMMARY #length 571 #molecular_weight 63638 #checksum 73

Query Match 43.7%; Score 1415; DB 2; Length 571;
 Best Local Similarity 42.0%; Pred. No. 3,52e-235;
 Matches 192; Conservative 104; Mismatches 160; Indels 1; Gaps 1;

DB 109 INDKIYSLANTNELEVLANKGETIENFVKEGVKADKFIYERKKKNTTTPVDISIDS 168
 6 VNDPILANNYDKKLLTHOGESIENRFKXEGNOLPDEFVYXERKKRSLSTNTSDIXVAT 65
 DB 169 VTDSTYAAIALANKGETENKPDPAVVTKRNPQKIHIDLPFGMDKAT-VEVNDPITYANVT 227
 66 NDSRLYFGALLVDEFLXENNPFLAVDAPMTYSIDLPGLASSDSFLOVEDPSNSVVG 125
 DB 228 AIDNLVMMQHDNYSGGVTLTPARTQYTESMYYSQIOEALANVASKILDGTLGIDFISISK 287

QY 126 AYNDLAKKHODYGVANNVPARKQYEKITAHSMEOLKVFSGDEFKTSGLDIDFNSVHS 185
 DB 288 GEKKVMIAAYKQIYTVSANLPNNPADVEDKSYTFEELORKGVSNAPPLMYSNVAAGRT 347
 186 GEKQIOIVNKKQIYTVSVDAVKNPGDVFODTYVEDLKORGISAEPLVYISXVAYGQ 245
 DB 348 VFVYKLETSKSNDEAFAALGTDVKTNGKSIDLSESTFAVLGGDAAEHKNVYTK 407
 246 VYKLETSKSNDEAFAALGTDVKTNGKSIDLSESTFAVLGGDAAEHKNVYTK 305
 DB 408 DFDIRNVIDKNAFTFSKKNPAPVISTYSVFLKKNKLAGVNNREYVETSTETYSKITL 467
 306 KVDMEVDLQESRFTADHPGLPISYTSFLRDNVAVAFQNSTDYVERKVTYVRNGD 365
 DB 468 SHQAYVAOYEILIMDEINDDKREYITRRDNMYSKTSFSTYVPLGANSRIRIMA 527
 366 DHSQAYVAOYIITWMELESTDHQKEVLTPKAWDRNGODLTAFTTISIPLKGNVRLSVK 425
 DB 528 RECTGLAEMWRTVYDEVYVPLASDINVSIMGTTLYP 564
 426 RECTGLAEMWRTVYEKTDLPVYRKRTISIMGTTLYP 462

RESULT 9
 ENTRY 139863 #type fragment
 TITLE hemolysin - Bacillus cereus (fragment)
 ORGANISM #formal name Bacillus cereus
 DATE 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996

ACCESSIONS 139863
 REFERENCE 139863
 #authors Yutsudo, T.; Okumura, K.; Iwasaki, M.; Hara, A.; Kakitani, S.; Mhamide, W.; Igarashi, H.; Hinuma, Y.
 #journal Infect. Immun. (1994) 62:4000-4004
 #title The gene encoding a new mitogenic factor in a Streptococcus pyogenes strain is a distributed only in group A streptococci.

#cross-references M01D:94341910
 #accession 139863
 #status preliminary; translated from GB/EMBL/DBJ
 #molecule_type DNA
 #residues 1-485 #label RES
 #cross-references GB:D21270; NID:9418066; PID:9600252
 SUMMARY #length 485 #checksum 3782

Query Match 41.2%; Score 1336; DB 2; Length 485;
 Best Local Similarity 39.8%; Pred. No. 4,43e-220;
 Matches 173; Conservative 117; Mismatches 144; Indels 1; Gaps 1;

DB 44 SIDTIGNLITVNOEVLAVNGDKVSEFVKESINSNGKFFVYVDRKHLQRHOSIFRLD 103
 5 AYNDPILANNYDKKLLTHOGESIENRFKXEGNOLPDEFVYXERKKRSLSTNTSDIXVAT 64
 DB 104 SVANRTYPGAVALANKAFADNOPSLVAKRKPPLNISTIDLPGRKENT-ITVONPTYGVA 162
 65 TINDSRLYFGALLVDEFLXENNPFLAVDAPMTYSIDLPGLASSDSFLOVEDPSNSVVR 124
 DB 163 GAVVDIVSTWNEKYSATHTLPARMQYTESMYYSKQAIASALNVAKYLDNSIDFNABA 222
 125 GAVNDLAKKHODYGVANNVPARKQYEKITAHSMEOLKVFSGDEFKTSGLDIDFNSVH 184
 DB 222 NEEKVYMAAKQIYTVSAELPNNPSDLFNSVYTFGLTIKGVSNAPPLMYSNVAAGR 282
 185 SEKQIOIVNKKQIYTVSVDAVKNPGDVFODTYVEDLKORGISAEPLVYISXVAYGR 244
 DB 283 IYVYKLETSKSKDVOAFAKLLKNNSVETSGQYKDIFFESTFTAVVLGGDAEHNKVVY 342
 245 GYKLETSKSNDEAFAALGTDVKTNGKSIDLSESTFAVLGGDAAEHKNVYTK 304
 DB 343 KDFNIRNITIDNALSFKKNPAPVISTYSVFLKDNATAAVHNTDYETTTETYSKAKT 402
 305 GRVMDVEDLQESRFTADHPGLPISYTSFLRDNVAVAFQNSTDYVERKVTYVRNGD 364

Db	403	LDHGAIVAAQEDVSNMGFEITDQNGKELTHKHWESGKDKTAHYSVLPDPNSKNIIV	465	2
Qy	365	LDHSGAIVAAQIYITLWELSLIDHOGKEVLIPKADRNGQDLTAHFTTSLPKGNRNLSVK	424	
Db	463	ARECTGLAEMWMRTI	477	
Qy	425	IRECTGLAEMWMRTV	439	
RESULT	10			
ENTRY		S47298	#type complete	
TITLE		sul1ysin - Streptococcus suis		
ORGANISM		#formal_name Streptococcus suis		
DATE		06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997		
ACCESSIONS		S47298		
REFERENCE		S47297		
#authors		Segers, R.		
#submitted		submitted to the EMBL Data Library, September 1994		
#accession		S47298		
#status		preliminary		
##molecule_type		DNA		
##residues		1-112 ##label SEG		
##cross-references		EMBL:J263608; NID:g535307; PID:g535308		
SUMMARY		#length 112 #molecular_weight 12464 #checksum 9264		
Query Match		14.3%; Score 464; DB 2; Length 112;		
Best Local Similarity		50.9%; Pred. No. 2,01e-57;		
Matches		57; Conservative 22; Mismatches 33; Indels 0; Gaps		
Db	1	PGVPYSISTVTXXXXXPAQXISNSEYIEFTSYHNSALTLKHSAGVAKYNITWEVSY	60	
Qy	325	PGLPISYTSISFLRNVVAFTQNSIDYETVYKVLAYRNGDLLDHSAGVAAQIYITWELSY	384	
Db	61	NEAGEEVEEXKAMPDKNGVNLSSHSEYIIXIPGNAXNLNVNIOECTGLAEMW	112	
Qy	385	DHOGKEVLIPKADRNGQDLTAHFTTSLPKGNRNLSVKIRECTGLAEMWM	436	
RESULT	11			
ENTRY		S47297	#type complete	
TITLE		sul1ysin - Streptococcus suis		
ORGANISM		#formal_name Streptococcus suis		
DATE		06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997		
ACCESSIONS		S47297		
REFERENCE		S47297		
#authors		Segers, R.		
#submitted		submitted to the EMBL Data Library, September 1994		
#accession		S47297		
#status		preliminary		
##molecule_type		DNA		
##residues		1-96 ##label SEG		
##cross-references		EMBL:J263607; NID:g535305; PID:g535306		
SUMMARY		#length 96 #molecular_weight 10471 #checksum 6662		
Query Match		6.6%; Score 214; DB 2; Length 96;		
Best Local Similarity		33.7%; Pred. No. 3,17e-15;		
Matches		30; Conservative 22; Mismatches 36; Indels 0; Gaps		
Db	1	KODNOYFOSLYTGPQELITNEGEYIDNPATGTGLENGFVVLRRKXITTDNSADITYI	60	
Qy	4	KAVNDFTILAMNDKRLTLTHQGESIENRFFXKEGNDLPDEFVYXERRKRSSTMTSDIXYI	63	
Db	61	XAKAANISPGALLRXKQNLNDNPILLI	89	
Qy	64	ATNDSRLYPGALLVVDFTLXENNPILLAV	92	
RESULT	12			
ENTRY		S03974	#type complete	
TITLE		amine oxidase (flavin-containing) (EC 1.4.3.4) A - bovine		
ALTERNATE_NAMES		monooamine oxidase type A		

```

ORGANISM      #formal_name Bos primigenius taurus #common_name cattle
DATE          28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change
              29-Jan-1999

ACCESSIONS    S03974
REFERENCE      #authors
               Powell, J.F.; Hsu, Y.P.P.; Weyler, W.; Chen, S.; Salach, J.;
               Andrikopoulos, K.; Maillet, J.; Breakfield, X.O.
               Biochem. J. (1988) 259:407-413
               The primary structure of bovine monamine oxidase type A.
               Comparison with peptide sequences of bovine monamine
               oxidase type B and other flavoenzymes.
               #cross-references MUID:89246344
               #accession      S03974
               #status         not compared with conceptual translation
               #molecule_type mRNA
               #residues       1-527 ##label POW
               #cross-references GB:X15609; NID:q523; PID:q524
               FAD: flavoprotein; mitochondrion; oxidoreductase

KEYWORDS      #region beta-alpha-beta FAD nucleotide-binding fold\
               #modified_site S-(8alpha-FAD)-cysteine (Cys) #status
               FEATURE         predicted
               15-43
               406

SUMMARY       #length 527 #molecular-weight 59800 #checksum 6598

Query Match   3.8%; Score 124; DB 2; Length 527;
Best Local Similarity 20.08; Pred. No. 1,66e-02;
Matches 30; Conservative 42; Mismatches 72; Indels 6; Gaps 5;

Db            1 MESLQKTSAGOMEDVVVIGGJISGLISAARK-DLAHEENVNVLAREHVGGRITYVNEH 59
               ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY            250 LETSKSAEVEAAFEALIKGVKVAPOTEKQILDMTEVKAVILGGDPSSGARVVTGKVDM 309
               ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

Db            60 V-DYVDGGAIVGPTQKRIIRLSKQGLETYVNNENRKYHYKQTPFGFAPPPWNP 118
               ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY            310 VEDLIQESRFTADHPGLPTSYTTF-LKDNVAATPQNSTDIYETKTATYKNGDLLIDHS 368
               ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

Db            119 IAYL-DYNNLMRTM--DNMGKEITPAAPME 145
               ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY            369 GAYVAQYITTMNELSTYHQCKEVLTPRKAMD 398
               ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

RESULT 13
ENTRY   S00552 #type complete
TITLE   mitochondrial processing peptidase (EC 3.4.99.41) beta chain
         precursor - yeast (Saccharomyces cerevisiae)
         mitochondrial processing proteinase 48K chain;
         processing-enhancing protein; protein L9632_10; protein
         YLR163c
ORGANISM #formal_name Saccharomyces cerevisiae
DATE      28-Feb-1990 #sequence_revision 14-Jul-1994 #text_change
              18-Sep-1998
ACCESSIONS S00552; A38734; S68479
REFERENCE   #authors
               Witte, C.; Jensen, R.E.; Yaffe, M.P.; Schatz, G.
               EMBO J. (1988) 7:1439-1447
               MASI, a gene essential for yeast mitochondrial assembly,
               encodes a subunit of the mitochondrial processing protease.
               #cross-references MUID:88312592
               #accession      S00552
               #molecule_type DNA
               ##residues      1-462 ##label WIT
               #cross-references EMBL:X07649; NID:q3886; PID:q3887
               A38734
REFERENCE   #authors
               Yang, M.; Gell, V.; Oppliger, W.; Suda, K.; James, P.;
               Schatz, G.
               J. Biol. Chem. (1991) 266:6416-6423
               The MAS-encoded processing protease of yeast mitochondria.
               Interaction of the purified enzyme with signal peptides and
               a purified precursor protein.
               #cross-references MUID:9117897
               #accession      A38734
               #molecule_type protein
               ##residues      21-32 ##label YAN

```

```

Best Local Similarity 18.8%; Pred. No. 4.51e-01;
Matches 33; Conservative 58; Mismatches 71; Indels 14; Gaps 13;

Db 567 SEEFTEEDDQFCIADIQVDSKISKSDTONPTNSIIDITSAASSIASEKCEIM 626
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 22 THQGS-LENEFXEGNOLDPEFVXERRKRSISTYTSIDIXVATGDSLTYPGAL-1IV 78
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 627 SOSKMEIROSILKTGLPMRTKYEIIISLOTASQIISTNPNDKGHHGVANF-SK-IEI 684
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 79 DELIKENNPITLAVDRAPMYTSIDL-PGLASDSFQIOWEDPSSSVKRGVNDLAKMHQD 137
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 138 YGVANN-VPARXQY-EKI-T-AH-SMEOLKVK-FGSD-FEKTGNSIDI-DENSVHS 185
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

RESULT 15
ENTRY 140468 #type complete
TITLE surface layer protein sbSa - Bacillus stearothermophilus
ORGANISM #formal_name Bacillus stearothermophilus
DATE 12-Aug-1996 #sequence-revision 12-Aug-1996 #text_change
18-Oct-1996
ACCESSIONS 140468; S34365
REFERENCE 140468
#authors Kuen, B.; Sleytr, U.B.; Lubitz, W.
#journal Gene (1994) 145:115-120
#title Sequence analysis of the sbSa gene encoding the 130-kDa
      surface-layer protein of Bacillus stearothermophilus strain
      PV72.

#cross-references MUID:94320770
#accession 140468
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
##residues 1-1228 ##label RSS
GENETICS #cross-references EMBL:X71092; NID:g312729; PID:g312730
#gene sbSa
SUMMARY #length 1228 #molecular-weight 131075 #checksum 2182

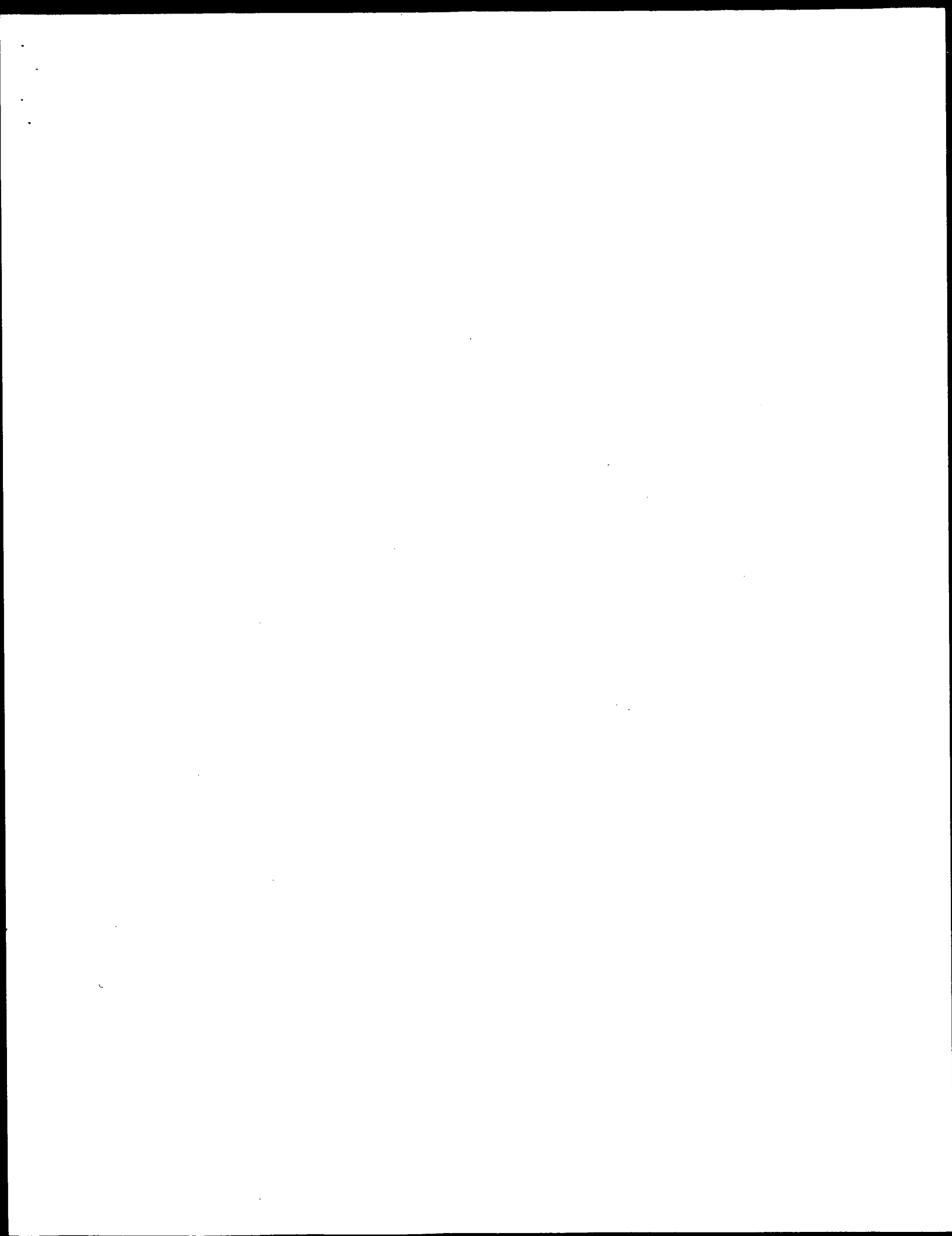
Query Match 3 5%; Score 114; DB 2; Length 1228;
Best Local Similarity 24.1%; Pred. No. 2.64e-01;
Matches 40; Conservative 44; Mismatches 72; Indels 10; Gaps 10;

Db 200 YDITV-AMKAR-EY-ODAVKAGNLDRAKAADVQIQYLPKYVDARFTELT-EVAKKALDA 255
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 200 YTVSVDAVDRKPGDVFQDITVEDLKRGISABRPVLYISXVAYGQVYLLKLETTSSKSEV 259
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 256 DEALTFRKVSYS-AINTQNKAVELTAYVNGT-LKLQLSAANDIYVNVTRIRKYGDG 313
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 260 -EAAFEALINGVAVAQTEWKQI-LDNTSVKAVILIGGPPSSGARVVTGKDVMEIDLIOEG 317
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

Db 314 NIPALNTADVSLSTDGKTIYDASPTPENNTEKYVWVGJKDKNG 359
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 318 S-RTADHPGLPISTYTSFLRDNVAVATPENSIDY-VETKIVAYNRNG 361
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

Search completed: Mon Aug 30 16:02:08 1999
Job time : 47 secs.

```



 W O R L D
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSPRCH_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 30 15:58:43 1999; Maspar time 14.67 seconds

Tabular output not generated. 907.846 Million cell updates/sec

Title: >US-09-120-044-3-COPY
 Description: (1-471) from dev1120044.pep
 Perfect Score: 3241
 Sequence: 1 MANKAVNDITLMMNDKKL.....TISGNTLTYPVEDKVENND 471

Scoring table: PAM 150
 Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot37
 1:swissprot

Statistics: Mean 51.515; Variance 97.371; scale 0.529

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3238	99.6	470	1	TACY_STRPN	0.00e+00
2	1599	49.3	500	1	TACY_CLOPE	0.00e+00
3	1582	47.6	529	1	TACY_LISTMO	1.20e-300
4	1523	47.0	528	1	TACY_LISTMO	2.01e-296
5	1476	45.5	530	1	TACY_LISTMO	5.56e-286
6	1468	45.3	501	1	TACY_LISTMO	3.32e-284
7	1415	43.7	571	1	TACY_STRPN	1.91e-272
8	1416	43.7	574	1	TACY_STRPN	1.14e-272
9	1389	43.2	574	1	TACY_STRPN	6.71e-269
10	1336	41.2	485	1	TACY_STRPN	5.24e-255
11	124	3.8	527	1	TACY_STRPN	2.99e-255
12	114	3.5	462	1	TACY_STRPN	5.17e-02
13	114	3.5	462	1	TACY_STRPN	5.17e-02
14	106	3.3	511	1	TACY_STRPN	5.56e-01
15	104	3.2	364	1	TACY_STRPN	9.87e-01
16	105	3.2	418	1	TACY_STRPN	7.42e-01
17	104	3.2	514	1	TACY_STRPN	9.87e-01
18	105	3.2	504	1	TACY_STRPN	7.42e-01
19	100	3.1	446	1	TACY_STRPN	3.02e+00
20	100	3.1	279	1	TACY_STRPN	3.02e+00
21	99	3.1	340	1	TACY_STRPN	3.98e+00
22	99	3.1	350	1	TACY_STRPN	3.98e+00
23	100	3.1	427	1	TACY_STRPN	3.02e+00

24	100	3.1	492	1	FLIC_SALRU	3.02e+00
25	100	3.1	502	1	SPIL_LYCES	3.02e+00
26	100	3.1	513	1	YVIL_ECOLI	3.02e+00
27	99	3.1	527	1	AMIE_OXIDSE	3.98e+00
28	101	3.1	722	1	Y442_YEAST	2.29e+00
29	102	3.1	756	1	Y328_YEAST	1.74e+00
30	99	3.1	878	1	SYV_METJA	3.98e+00
31	101	3.1	1123	1	V120_HSV1	2.29e+00
32	99	3.1	1137	1	MDP_BACSU	3.98e+00
33	100	3.1	1235	1	CY44_TRYBB	3.02e+00
34	99	3.1	1681	1	CLH_CAEEL	3.98e+00
35	100	3.1	1744	1	TANA_XENLA	3.02e+00
36	98	3.0	309	1	RAS1_YEAST	5.22e+00
37	97	3.0	325	1	SBOD_BACSU	6.83e+00
38	96	3.0	349	1	PHOL_MYCLE	8.91e+00
39	96	3.0	410	1	NEUS_MOUSE	8.91e+00
40	97	3.0	419	1	Y223_MYCPN	6.83e+00
41	97	3.0	667	1	GR78_APLCA	6.83e+00
42	97	3.0	770	1	A4_RAT	6.83e+00
43	97	3.0	770	1	A4_MOUSE	6.83e+00
44	97	3.0	1026	1	BGAL_STRTR	6.83e+00
45	98	3.0	1675	1	CLH_RAT	5.22e+00

ALIGNMENTS

RESULT 1
 ID TACY_STRPN STANDARD; PRT: 470 AA.
 AC P11990;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE PNEUMOLYSIN (THIOL-ACTIVATED CYTOLYSIN).
 GN PLY.
 OS STREPTOCOCCUS PNEUMONIAE.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
 OC STREPTOCOCCUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NTCC 7466 SEROTYPE II;
 RX MEDLINE; 87193109.
 RA WALKER J.A., ALLEN R.L., FALMAGNE P., JOHNSON M.K., BOULNOIS G.J.;
 RT "Molecular cloning, characterization, and complete nucleotide
 RT sequence of the gene for pneumolysin, the sulfhydryl-activated toxin
 RT of Streptococcus pneumoniae";
 RL INFECT. IMMUN. 55:1184-1189(1987).
 CC -!- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
 CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
 CC EUKARYOTIC CELL MEMBRANES.
 CC -!- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X52474; G47404; -.
 CC EMBL: M17717; G153692; -.
 CC PIR: A28568; A28568.
 CC PROSITE: PS00461; THIOL_CYTOLYSINS; 1.
 CC DR PFAM: PF01289; THIOL_CYTOLYSIN; 1.
 CC DR HSSP: P19995; IPEO.
 CC FT TOXIN; HEMOLYSIN; CYTOLYSIS.
 CC INIT MET 0
 CC ACT_SITE 427 427 BINDING TO CHOLESTEROL (BY SIMILARITY).
 CC SEQUENCE 470 AA; 52768 MW; D3F3A252 CRC32;
 CC
 CC Query Match 99.6%; Score 3238; DB 1; Length 470;
 CC Best Local Similarity 98.1%; Pred. No. 0.00e+00;

Matches 461: Conservative 1: Mismatches 8: Indels 0: Gaps 0:

```

Db 1 ANAVNDFIAMNYDKKLLTHOGESIEENFKIEGNOLEPDEFVIERKRSISTNTSDIS 60
   |||||
Qy 2 ANKAVNDFIAMNYDKKLLTHOGESIEENFKIEGNOLEPDEFVIERKRSISTNTSDIX 61
   |||||
Db 61 VTATNDSRLYPGALLVDETLLENPTLLAVDRAPMYISIDLPGLASSDSFQVEDPSSNS 120
   |||||
Qy 62 VTATNDSRLYPGALLVDETLLENPTLLAVDRAPMYISIDLPGLASSDSFQVEDPSSNS 121
   |||||
Db 121 SVGAANDLLAKKHODYGOVANNYPARMQYETKITAHSMEOLKAFGSDPFKTSNIDIDPN 180
   |||||
Qy 122 SVGAANDLLAKKHODYGOVANNYPARMQYETKITAHSMEOLKAFGSDPFKTSNIDIDPN 181
   |||||
Db 181 SVHSGEKQIOIVFKOITYTVSDAVKPNQDVQDTVTYVEDLKQRGISAREPLVYISSVA 240
   |||||
Qy 182 SVHSGEKQIOIVFKOITYTVSDAVKPNQDVQDTVTYVEDLKQRGISAREPLVYISSVA 241
   |||||
Db 241 YGRQVYLKLETTSKSDEVEAFALIKGVVAPQTEMKOILNTEVKAVILGDESSGAR 300
   |||||
Qy 242 YGRQVYLKLETTSKSDEVEAFALIKGVVAPQTEMKOILNTEVKAVILGDESSGAR 301
   |||||
Db 301 VVTGKVDWEDLLIOESRFTADHPGLPISTYTSFLRDNVATFONSTDVYETKVTAYRNG 360
   |||||
Qy 302 VVTGKVDWEDLLIOESRFTADHPGLPISTYTSFLRDNVATFONSTDVYETKVTAYRNG 361
   |||||
Db 361 DILLDSGAVVAOYITWDELSTDHOGKEVLLPKAMDNGODLJAHTTSIDPLKGNVRL 420
   |||||
Qy 362 DILLDSGAVVAOYITWDELSTDHOGKEVLLPKAMDNGODLJAHTTSIDPLKGNVRL 421
   |||||
Db 421 SVKIRCTGLAMEMWRTVEKTDLPYRKRTISINGTTLTYPOVEKVENMD 470
   |||||
Qy 422 SVKIRCTGLAMEMWRTVEKTDLPYRKRTISINGTTLTYPOVEKVENMD 471
   |||||

RESULT
ID 2 TACY CLOPE STANDARD: PRT: 500 AA.
AC 19995:
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE PERRINGOLYSIN O PRECURSOR (THETA-TOXIN) (THIOL-ACTIVATED CYTOLYSIN).
GN PFO OR PFOR OR PROA.
OS CLOSTRIDIUM PERFRINGENS.
OC PLASMID PRT1B.
OC BACTERIA: FIRNICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;
OC CLOSTRIDIUM.
OC [1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE: 9109951.
RA SHIMIZU T., OKABE A., MINAMI J., HAYASHI H.;
RT "An upstream regulatory sequence stimulates expression of the
RL perfringolysin O gene of Clostridium perfringens.";
RT INFECT. IMMUN. 59:137-142(1991).
RN RP [2]
RN RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-44.
RX MEDLINE: 89032623.
RA TWETEN R.K.;
RT "Nucleotide sequence of the gene for perfringolysin O (theta-toxin)
RL from Clostridium perfringens: significant homology with the genes for
RT streptolysin O and pneumolysin.";
RT INFECT. IMMUN. 56:3235-3240(1988).
RN RP [3]
RN RP SEQUENCE OF 29-45 AND 305-312.
RX STRAIN=PB6K;
RX MEDLINE: 87076517.
RA OHNO-IWASHITA Y., IWAMOTO M., MITSUI K., KAWASAKI H., ANDO S.;
RT "Cold-labile hemolysin produced by limited proteolysis of theta-toxin
RL from Clostridium perfringens.";
RT BIOCHEMISTRY 25:6048-6053(1986).
RN RP [4]
RN RP SEQUENCE OF 492-500 FROM N.A.

```

RC STRAIN=NCTC 8237;
RX MEDLINE: 96123363.
RA SHIMIZU T., KOBAYASHI T., BA-THIEN W., OHTANI K., HAYASHI H.;
RT "Sequence analysis of flanking regions of the pfoA gene of
RT Clostridium perfringens: beta-galactosidase gene (pbg) is located in
RL the 3'-flanking region.";
RT MICROBIOL. IMMUNOL. 39:677-686(1995).
RN RP [5]
RN RP IMPORTANCE OF THIOL-GROUP IN CHOLESTEROL BINDING.
RX MEDLINE: 88004463.
RA IWAMOTO M., OHNO-IWASHITA Y., ANDO S.;
RT "Role of the essential thiol group in the thiol-activated cytolysin
RT from Clostridium perfringens.";
RL EUR. J. BIOCHEM. 167:425-430(1987).
RN RP [6]
RN RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE: 97113505.
RA FELL S.C., ROSSJOHN J., ROHDE K., TWETEN R.K., PARKER M.W.;
RT "Crystallization and preliminary x-ray analysis of a thiol-activated
RL cytolysin.";
RL FEBS LETT. 397:290-292(1996).
RN RP [7]
RN RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE: 97325744.
RA ROSSJOHN J., FELL S.C., MCKINSTRY W.J., TWETEN R.K., PARKER M.W.;
RT "Structure of a cholesterol-binding, thiol-activated cytolysin and a
RT model of its membrane form.";
RL CELL. 89:685-692(1997).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SUBUNIT: FORMS OLIGOMERS IN THE HOST MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M61080; G144886; -;
DR EMBL: M36704; G144884; -;
DR EMBL: D49537; G1502275; -;
DR PIR: B43577; B43577.
DR PDB: 1PRO: 05-AUG-98.
DR PROSITE: P500481; THIOL_CYTOLYSINS; 1.
DR PFAM: PFO1289; Thiol_cytolysin; 1.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMID; 3D-STRUCTURE.
FT SIGNAL 1 28
FT CHAIN 29 500
FT ACT SITE 459 459
FT CONFLICT 30 30
FT CONFLICT 34 34
FT CONFLICT 126 128
SQ SEQUENCE 500 AA: 55799 MW; 17FC32BC CRC32;
Query Match 49.3%; Score 1599; DB 1; Length 500;
Best Local Similarity 46.5%; Pred. No. 0.00e+00;
Matches 214: Conservative 105: Mismatches 140: Indels 1; Gaps 1;

```

Db 35 NOSIDGSSLSYNRNEVLASNGDKIESFVPEKGGKAGKFIYEROKRSITTSVPDISI 94
   |||
Qy 3 NKAIVNFIAMNYDKKLLTHOGESIEENFKIEGNOLEPDEFVIERKRSISTNTSDIX 62
   |||
Db 95 IDSVDNRYTPGALQADKAIVENPTLLVKKRPININIDLPGLGKENS-IKVDPPYK 153
   |||
Qy 63 TATNDSRLYPGALLVDETLLENPTLLAVDRAPMYISIDLPGLASSDSFQVEDPSSNS 122
   |||
Db 154 VSGAIDELVSKNNEKSSHTLTPARTQYSESVMYKSOISLANNVAVLNSLGVDINA 213
   |||

```



```

GN ILO.
OC LISTERIA IVANOVII.
OC LISTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19119;
RC MEDLINE; 92182018.
KA HAAS A., DUMESKY M., KREFT J.;
RT "Listeriolysin genes: complete sequence of ilo from Listeria ivanovii
and of ilo from Listeria seeligeri.";
RL BIOCHIM. BIOPHYS. ACTA 1130:81-84(1992).
CC -I- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
EUKARYOTIC CELL MEMBRANES.
CC -I- SIMILARITY: BELONGS TO THE THIOI-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60461; E39033; ALU.INT.
DR PROSITE; PS00481; THIOI_CYTOLYSINS; 1.
DR PFAM; PF01289; Thiol_cytolysin; 1.
DR HSP; P19995; IPEO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 23
FT CHAIN 24 528
FT ACT_SITE 483 483 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 528 AA; 58542 MW; CFAE84AD CRC32;

Query Match 47.0%; Score 1523; DB 1; Length 528;
Best Local Similarity 42.7%; Pred. No. 2,01e-296;
Matches 199; Conservative 119; Mismatches 147; Indels 1; Gaps 1;

DB 60 IDOYIOGLDYKNNILVYDGEAVKVPKAGYKEGNOYIYVEKKKRSINONNADIYVINS 119
OY 6 VNDFLANNYKRLTLHQGESIENRFKKEGNOPLDEFVYXERKRRLSTNTSDIXYAT 65
DB 120 LASLTYPGALVANSSELVENOPVLPVKRDSVTLSDLPKGVNHDNIYVQMATKSNIND 179
OY 66 NDSRLYPGALVVDLTKENNPILLAVDRAPIYSIDLPGIASSDSFLQVEDPSNSVVG 125
DB 180 GVNTLVDRMNNKYSSEYEPNISAKIDYQEMAYSESOLVAFKGAFAKVNNSLVNFEAIS 239
OY 126 AVNDLAKMHDDYGVN-NVPAKQYEKITHASMEOLKVFSGDFEKTGNSLIDEFNSVH 184
DB 240 EGVQVEEYINFKQIYYIVNNEPTSPKFEKGSYTKENLOALGVNAENPAYISSVAYGR 299
OY 185 SGEKQIOIVNKKQIYYIVSVDAVKNPGDFQDVTYVEDLKORGISARPLVYISXVAYGR 244
DB 300 DIFKYLSSSSSTVYKAFDTAFKGSYKGDTELENTIIONAFAVYIGGSANDEVEIID 359
OY 245 QYVTLKLETTSSKXEYEAFAELIKGVKAVAPQTEWKQILDNTEVAVYLLGDPSSGARVYT 304
DB 360 GDLKLRDLKOGANFPKKNKNGVPPIATNTNLKNDLAVVKNNSSEYLETTSKAYSQKIN 419
OY 305 GKVDMVEDLIDEGSRFTADHPGLPISITTSFLRNVAVATQNSDYEIVETVATYARNDDL 364
DB 420 LDHSGAVARENVWDESVYDANGNEVEYKNSKENDKLAHFTTSIYPGNARNINIH 479
OY 365 LDHSGAVAYQIYIWNELSYDHQGEVLTLPKAMRNGODLTAHFTTISPLKGVNRLSVK 424
DB 480 AKECTGLAMEWRTVVDRLPLVKNRNVCIWGTITLYPAVSQVYDN 525
OY 425 IRECTGLAMEWRTVYERKTDPLVRKRRTISIMGTTLYPQVEDKVEN 470

```

```

RESULT 5
ID TACY LISSE STANDARD; PRT; 530 AA.
AC P31830;
DT 01-JUL-1993 (REL. 26 CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE SEELIGERIOLYSIN PRECURSOR (THIOI-ACTIVATED CYTOLYSIN).
GN LSO.
OS LISTERIA SEELIGERI.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC LISTERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SLCC;
RC MEDLINE; 92182018.
KA HAAS A., DUMESKY M., KREFT J.;
RT "Listeriolysin genes: complete sequence of ilo from Listeria ivanovii
and of ilo from Listeria seeligeri.";
RL BIOCHIM. BIOPHYS. ACTA 1130:81-84(1992).
CC -I- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
EUKARYOTIC CELL MEMBRANES.
CC -I- SIMILARITY: BELONGS TO THE THIOI-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60462; G44145; -.
DR PIR; S22340; S22340.
DR PROSITE; PS00481; THIOI_CYTOLYSINS; 1.
DR PFAM; PF01289; Thiol_cytolysin; 1.
DR HSP; P19995; IPEO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 530
FT ACT_SITE 485 485 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 530 AA; 59181 MW; 66A497A5 CRC32;

Query Match 45.5%; Score 1476; DB 1; Length 530;
Best Local Similarity 42.5%; Pred. No. 5.56e-286;
Matches 198; Conservative 111; Mismatches 156; Indels 1; Gaps 1;

DB 62 IKTYTGLNIDKNSILVYQGAVTNVPPKGYKDGSEYIYVEKKKRSINONNADIVINA 121
OY 6 VNDFLANNYKRLTLHQGESIENRFKKEGNOPLDEFVYXERKRRLSTNTSDIXYAT 65
DB 122 ISLTYPGALVANSSELVENOPVLPVKRDSVTLSDLPKGVNHDNIYVQMATKSNIND 179
OY 126 AVNDLAKMHDDYGVN-NVPAKQYEKITHASMEOLKVFSGDFEKTGNSLIDEFNSVH 184
DB 182 AVNTLVDRMNNKYSKAYEPNISAKIDYQEMAYSESOLIAKFGAFKVNNSLVNFEAIS 241
OY 66 NDSRLYPGALVVDLTKENNPILLAVDRAPIYSIDLPGIASSDSFLQVEDPSNSVVG 125
DB 240 EGVQVEEYINFKQIYYIVNNEPTSPKFEKGSYTKENLOALGVNAENPAYISSVAYGR 299
OY 185 SGEKQIOIVNKKQIYYIVSVDAVKNPGDFQDVTYVEDLKORGISARPLVYISXVAYGR 244
DB 300 DIFKYLSSSSSTVYKAFDTAFKGSYKGDTELENTIIONAFAVYIGGSANDEVEIID 359
OY 245 QYVTLKLETTSSKXEYEAFAELIKGVKAVAPQTEWKQILDNTEVAVYLLGDPSSGARVYT 304
DB 360 GDLKLRDLKOGANFPKKNKNGVPPIATNTNLKNDLAVVKNNSSEYLETTSKAYSQKIN 419
OY 305 GKVDMVEDLIDEGSRFTADHPGLPISITTSFLRNVAVATQNSDYEIVETVATYARNDDL 364
DB 420 LDHSGAVARENVWDESVYDANGNEVEYKNSKENDKLAHFTTSIYPGNARNINIH 479
OY 365 LDHSGAVAYQIYIWNELSYDHQGEVLTLPKAMRNGODLTAHFTTISPLKGVNRLSVK 424
DB 480 AKECTGLAMEWRTVVDRLPLVKNRNVCIWGTITLYPAVSQVYDN 525
OY 425 IRECTGLAMEWRTVYERKTDPLVRKRRTISIMGTTLYPQVEDKVEN 470
DB 422 IDHSGGIYVQAINISWDESVYDANGNEIKVHKWGENYKSLAHFTTSIYPGNARNINIH 481

```



```

QY 365 LDHSGAVYAOYITWMLSDYDHGKEVLTPKAMDNRGODLTAHFTSIPLKGAVRNLSYK 424
DB 482 ARDTGTFEMWRTYIDRLNPLVKNNSVINGTTLTPKRSNVDN 527
QY 425 IRECTGLAMWMTVEKTDLPVLRKRTISIMGTITLPVEDAKVEN 470

RESULT 6
ID TACY PAEAL STANDARD; PRT: 501 AA.
AC P23564;
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ALVEOLYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN ALV.
OS PARENTHACILLUS ALVEI (BACILLUS ALVEI).
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
RN PAENIBACILLUS.
RP SEQUENCE FROM N.A.
RA MEDLINE; 91072294.
RX GEOFFROY C., MENGAUD J., ALOUF J.E., COSSART P.;
RT "Alveolysin, the thiol-activated toxin of Bacillus alvei, is
RT homologous to listeriolysin O, pertingolysin O, pneumolysin, and
RT streptolysin O and contains a single cysteine."
RT J. BACTERIOL. 172:7301-7305(1990).
RN [2]
RP SEQUENCE OF 33-60.
RA ALOUF J.E., GEOFFROY C., GILLES A.M., FALMAGNE P.;
RT (IN) RAPPUOLI R., ALOUF J.E., FALMAGNE P. (EDS.);
RT BACTERIAL PROTEIN TOXINS, PP.49-50, GUSTAV FISHER VERLAG,
RT STUTTGART (1990).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M62709; G142473; -.
DR PIR; A37858; A37858.
DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM; PF01289; Thiol_cytolysin; 1.
DR HSSP; P19995; IPRO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 32
FT CHAIN 33 501
FT ACT_SITE 461 461 ALVEOLYSIN.
FT CONFLICT 56 56 BINDING TO CHOLESTEROL (BY SIMILARITY).
FT SEQUENCE 501 AA; 35268 MW; 8DF2C94A CRC32;

Query Match 45.3%; Score 1468; DB 1; Length 501;
Best Local Similarity 41.9%; Pred. No. 3,32e-284;
Matches 190; Conservative 121; Mismatches 141; Indels 1; Gaps 1;

DB 44 TAGLNYNNEVLAIQGOISSFVPEKGIQNGKFTIVERDKSLTSPVDISIVDSITNR 103
QY 10 ILAMVYDKKLLTGESIEENRFKEGNQDLPDEFVYXERKRKSLSTJMSDIXVATNDSR 69
DB 104 TYPGAIQLANDFADNOSLYMAARKPLDISLPGLKNEVT-ISYQNPNGYSSAIDQ 162
QY 70 LYPGALLVVDDELTKENNPLLAVDAPMTATYSIDLPGLASSDSFQVEDPSSSVGAVND 129
DB 163 LVSTGKYSHTLTPARLQYAESWYSONQISSALNANAVLNGTIGIDNNAVANEKK 222

```

```

QY 130 LIAKHODYGVQNNVPARKQYKAITAHSMEOIKYKFGSDEKTSNIDIDFNSYSGEKQ 189
DB 223 VVVAAYKQIEFTVSAGLPNNPSDLFSDSVFAELARGVSEAPPLWNVAVAGRTIYK 282
QY 190 IQIVNKKQIYTVSDAVKANPGDVFQDVIYVEDLKQKISAEPLVITISXVAGROYLK 249
DB 283 LETTSKNDVOTAFKLLNNPSIOASGOYKDIYENSFTAVLGGDQOTHNQVTKDFNV 342
QY 250 LETTSKSKVEEAFALLKGVKVAPOFEWKOILDNTEVKAVILGGDSSGARVYTGAVDM 309
QY 343 IOSVIKDNAOFSSKNPAPVISTYSVFLKDNISAAVHNTEYIEFTKTEYSKGIKLDHSG 402
QY 310 VEDLIOESRFTADHPGDPISYTSFTSLRDNVAVATFONSTDEYETKVAYRNGDILLDHS 369
DB 403 AVVAQFEVYVMEFSVDAGCEIVTRKSWDGMWRBSAHFSEITLPPNANKIRIFARECT 462
QY 370 AVVAQYITWMLSDYDHGKEVLTPKAMDNRGODLTAHFTSIPLKGAVRNLSVIRECT 429
DB 463 GLAMWMTVEKTDLPVLRKRTISIMGTITLP 495
QY 430 GLAMWMTVEKTDLPVLRKRTISIMGTITLP 462

RESULT 7
ID TACY STRPY STANDARD; PRT: 571 AA.
AC P21131;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS PYOGENES.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE; 88057628.
RX KEHOE M.A., MILLER L., WALKER J.A., BOUINIS G.J.;
RT "Nucleotide sequence of the streptolysin O (SLO) gene: structural
RT homologies between SLO and other membrane-damaging, thiol-activated
RT toxins."
RT INFECT. IMMUN. 55:3228-3232(1987).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M18638; G153811; -.
DR PIR; A43507; A43507.
DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM; PF01289; Thiol_cytolysin; 1.
DR HSSP; P19995; IPRO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMID.
FT SIGNAL 1 33
FT CHAIN 34 571
FT ACT_SITE 530 530 STREPTOLYSIN O.
FT SEQUENCE 571 AA; 63638 MW; 33124E54 CRC32;

Query Match 43.7%; Score 1415; DB 1; Length 571;
Best Local Similarity 42.0%; Pred. No. 1.91e-272;
Matches 192; Conservative 104; Mismatches 160; Indels 1; Gaps 1;

DB 109 INDKYISLWNEVLAKNGEITENFVPEKGVKADKRTIVIERKKKNNTPTVDISIDS 168

```

```

QY 6 VNDPFLAAMVYDKKLLTHOGESIEENRFKEGQJLPEDEVVEXERRKRSJSTNTSDIXVYAT 65
DB 169 VYDRTYPAALQIANKGFTENKPDVAVVTKRNPKIHDLPKMGDKAT-VEVNDPTVAVNT 227
QY 66 NDSRLYVPGALLVVDDELTKENNPTLLAVDRAPMVYSIDLPGIASSPSFLQVDDPNSVYRG 125
DB 228 AIDNVLNOMHNDVSGGNTLPARTQVTESMVYSKQIOEALVNVNSKIIDGLTIGDFKSIK 287
QY 126 AVNDLLAMWHDYGVGVNVNVPARKQYKJTAHSMQJLKFGSDFEKTNISLIDIFNSVHS 185
DB 288 GEKQVIAAYKQIEFTVSANLPNNPADYEDKSVTLKELORRGVSEAPPLFVSNVAVGRT 347
QY 166 GEKQIYVXKQIYTVSDAVKNGDVFQDTVYEDLKQIGISAERPLVYSXVAVGRQ 245
DB 348 VYVKLETSKSDVDEAFSAALKGTGVKTNKYSIDLENSFTAVVLGGDAEHNKVVTK 407
QY 246 VYKLETSKSDVDEAFSAALKGTGVKTNKYSIDLENSFTAVVLGGDAEHNKVVTK 305
DB 408 DEDVIRNVYKDNATSRKNPAPVISTSVFLKNNKIAGVNNREYVETTSYTGKINTL 467
QY 306 KYDWEDEDIQESSRFTADHPGLPISTYTSFLRDVNVAFQNSTDYETVKVAVYANGDLL 365
DB 468 SHOGAYVAYQYELIMDEINVDKGEVITKRRDNNWYSKTSPESTVPLIGANSNIRIMA 527
QY 366 DHSGAYVAYQYITWNELSYDHQGEVITLPAKMDRNGODLTHFTTSIPLGAVNANISVKI 425
DB 528 RECTGLAEMWVRKYVDEEDVYKLSKEINVINISGTLSP 564
QY 426 RECTGLAEMWVRKYVDEEDVYKLSKEINVINISGTLSP 462

```

```

RESULT 8
ID TACY STREQ STANDARD; PRT: 574 AA.
AC Q54114;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SIO.
OS STREPTOCOCCUS EOUISIMILIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-STMD-1;
RX MEDLINE; 95102113.
RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
RA IGARASHI H., YUTSUDO T.;
RT "Cloning and sequencing the streptolysin O genes of group C and group
RT G streptococci."
RL DNA SEQ. 4:325-328(1994).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: D16824; G498301;
CC PROSITE: PS00481; THIOI. CYTOLYSINS; 1.
CC DR PFAM: PF01289; Thiol. cytolysins; 1.
CC DR HSPR: P19995; 1PFO.
CC KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
CC FT SIGNAL 1 36 BY SIMILARITY.
CC CHAIN 37 574 STREPTOLYSIN O.

```

```

FT ACT_SITE 533 533 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 574 AA; 63991 MW; 98BF31B4 CRC32;
Query Match 43.7%; Score 1416; DB 1; Length 574;
Best Local Similarity 41.8%; Pred. No. 1,14e-272;
Matches 191; Conservative 106; Mismatches 159; Indels 1; Gaps 1;
DB 112 INDKYISLVNLEVLANKGFTIENFVREGVKKADKRTVIERKKKNTNTPVDISIDS 171
QY 6 VNDPFLAAMVYDKKLLTHOGESIEENRFKEGQJLPEDEVVEXERRKRSJSTNTSDIXVYAT 65
DB 172 VYDRTYPAALQIANKGFTENKPDVAVVTKRNPKIHDLPKMGDKAT-VEVNDPTVAVNT 230
QY 66 NDSRLYVPGALLVVDDELTKENNPTLLAVDRAPMVYSIDLPGIASSPSFLQVDDPNSVYRG 125
DB 231 AIDNVLNOMHNDVSGGNTLPARTQVTESMVYSKQIOEALVNVNSKIIDGLTIGDFKSIK 290
QY 126 AVNDLLAMWHDYGVGVNVNVPARKQYKJTAHSMQJLKFGSDFEKTNISLIDIFNSVHS 185
DB 291 GEKQVIAAYKQIEFTVSANLPNNPADYEDKSVTLKELORRGVSEAPPLFVSNVAVGRT 350
QY 186 GEKQIYVXKQIYTVSDAVKNGDVFQDTVYEDLKQIGISAERPLVYSXVAVGRQ 245
DB 351 VYVKLETSKSDVDEAFSAALKGTGVKTNKYSIDLENSFTAVVLGGDAEHNKVVTK 410
QY 246 VYKLETSKSDVDEAFSAALKGTGVKTNKYSIDLENSFTAVVLGGDAEHNKVVTK 305
DB 411 DEDVIRNVYKDNATSRKNPAPVISTSVFLKNNKIAGVNNREYVETTSYTGKINTL 470
QY 306 KYDWEDEDIQESSRFTADHPGLPISTYTSFLRDVNVAFQNSTDYETVKVAVYANGDLL 365
DB 471 SHOGAYVAYQYELIMDEINVDKGEVITKRRDNNWYSKTSPESTVPLIGANSNIRIMA 530
QY 366 DHSGAYVAYQYITWNELSYDHQGEVITLPAKMDRNGODLTHFTTSIPLGAVNANISVKI 425
DB 531 RECTGLAEMWVRKYVDEEDVYKLSKEINVINISGTLSP 567
QY 426 RECTGLAEMWVRKYVDEEDVYKLSKEINVINISGTLSP 462

```

```

RESULT 9
ID TACY STREQ STANDARD; PRT: 574 AA.
AC Q53957;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SIO.
OS STREPTOCOCCUS CANIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 95102113.
RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
RA IGARASHI H., YUTSUDO T.;
RT "Cloning and sequencing the streptolysin O genes of group C and group
RT G streptococci."
RL DNA SEQ. 4:325-328(1994).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----

```



```
RT      "NOP56 is a small nucleolar ribonucleoprotein component required for  
RL pre-18S rRNA processing in yeast.";  
CC J. BIOL. CHEM. 273:16453-16463(1998).  
CC -1- FUNCTION: REQUIRED FOR PRE-18S RNA PROCESSING.  
CC -1- SUBCELLULAR LOCATION: NUCLEUS; NUCLEOLUS.  
CC -1- SIMILARITY: BELONGS TO THE NOP5/NOP56 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
FT or send an email to license@isb-sib.ch).  
-----  
DR EMBL; X90565; G940841; -.  
DR EMBL; Z75217; E252143; -.  
DR EMBL; AF056070; G3170533; -.  
DR SGD; L000400; NOP5.  
KW NUCLEAR PROTEIN; RRNA PROCESSING.  
FT DOMAIN 441 511 ASP/GLU/LYS-RICH.  
SO SEQUENCE 511 AA; 56956 MW; F0A4F72C CRC32;  
  
Query Match 3.3%; Score 106; DB 1; Length 511;  
Best Local Similarity 22.1%; Pred. No. 5,566-01;  
Matches 30; Conservative 39; Mismatches 59; Indels 8; Gaps 8;  
  
Db 188 YGWH-FPELATIV-IDSVAVARILITMGIRKASSTDSLPEIEER-VKTALEVSMG 244  
QY |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
242 YGRVYYLTLETTSKXEAAPFAIKVKV-APOETEKOLL-DNTEVKAVALIGSDPSSG 299  
DB 245 TEITOTDIDNINALAEQIJEPAAYEEQS-NLSARMKAIPNLTOVGELGARLIA-H 302  
QY :||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
300 ARVYAGKVDWEDDLQESRFTADHPGLISTYSFLDNVATFNQNSTDYEVETAYR 359  
Db 303 SGLSLAKSPASTTQ 318  
QY :||:|||||:  
QY 360 NGDDL-LDHSGAVYAQ 374  
  
RESULT 15  
ID GCH2_PROBLE STANDARD; PRT; 364 AA.  
AC Q02008;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE GTP CYCLOHYDROLASE II (EC 3.5.4.25) / 3,4-DIHYDROXY-2-BUTANONE 4-  
DE PHOSPHATE SYNTHASE (DHBP SYNTHASE).  
GN RIBA.  
OS PHOTOBACTERIUM LEIOGNATHI.  
OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; VIBRIONACEAE;  
OC PHOTOBACTERIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 25521;  
RX MEDLINE; 92360014.  
RA LEE C.Y., MEIGHEN E.A.;  
RT "The lux genes in Photobacterium leioognathi are closely linked with  
RT genes corresponding in sequence to riboflavin synthesis genes";  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 186:690-697(1992).  
RL -1- CATALYTIC ACTIVITY: GTP + 3 H(2O) = FORMATE + 2,5-DIMINO-6-  
CC HYDROXY-4-(5-PHOSPHORIBOSYLIAMINO)PYRIMIDINE + PYRROPHOSPHATE.  
CC -1- PATHWAY: RIBOFOLPIN BIOSYNTHESIS.  
CC -1- SIMILARITY: TO OTHER BACTERIAL GTP CYCLOHYDROLASE II (RIBA), AND  
CC TO E.COLI RIBB AND V.HARVEYI LOXH.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).
```

 (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mprch_dp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 30 15:59:45 1999; MasPar time 29.02 Seconds

Tabular output not generated. 885.876 Million cell updates/sec

Title: >US-09-120-044-3-COPY
 Description: (1-471) from dev1120044.pep
 Perfect Score: 3241
 Sequence: 1 MANKAVNDFILAMNDKKL.....TISTGTTLYPQVEDKYEND 471

Scoring table: PAM 150
 Gap 11

Searched: 179066 segs, 54579741 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

sptrembl9
 1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
 5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
 9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
 13:sp_vertebrate 14:sp_virus

Statistics: Mean 49.380; Variance 93.387; scale 0.529

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	1779	54.9	497	2	055996	SULIYSIN.	0.00e+00
2	1774	54.7	497	2	085102	HEMOLYSIN.	0.00e+00
3	1302	40.2	534	2	031241	POLYLYSIN.	1.55e-248
4	124	3.8	50	2	048772	LISTERIOLYSIN O (FRAGM	1.79e-03
5	118	3.6	50	2	048773	LISTERIOLYSIN O (FRAGM	1.75e-02
6	117	3.6	450	2	068518	HYPOHETICAL 49.7 KD P	8.43e-02
7	112	3.5	748	3	012098	CHROMOSOME XII READING	6.18e-02
8	113	3.5	1361	2	087083	S-LAYER PROTEIN.	6.18e-02
9	113	3.5	1361	2	030524	S-LAYER PROTEIN.	6.18e-02
10	111	3.4	187	1	058072	187AA LONG HYPOHETICA	1.15e-01
11	111	3.4	814	10	085715	HYPOHETICAL 91.9 KD P	1.15e-01
12	107	3.3	218	1	058691	HYPOHETICAL PROTEIN M	3.89e-01
13	106	3.3	265	3	012044	CHROMOSOME XV READING	5.25e-01
14	106	3.3	501	1	028947	CONSERVED HYPOHETICAL	3.89e-01
15	107	3.3	710	2	085906	TRANSPONSON PROTEIN B.	3.89e-01
16	107	3.3	5027	11	063460	PROLINE-RICH PROTEIN (3.89e-01
17	103	3.2	303	3	005541	CHROMOSOME IV COSMID 9	1.27e+00
18	105	3.2	426	2	067578	FLAGELLAR HOOK ASSOCI	7.07e-01
19	103	3.2	640	5	009647	HYPOHETICAL 74.1 KD P	1.27e+00
20	104	3.2	867	2	068185	CLPB CHAPERONE HOMOLOG	9.51e-01

RESULT ID	1	PRELIMINARY:	PRT:	497 AA.
AC	055996	055997:		
DT	01-NOV-1996	(TREMBLREL. 01, CREATED)		
DT	01-JAN-1999	(TREMBLREL. 09, LAST SEQUENCE UPDATE)		
DT	01-JAN-1999	(TREMBLREL. 09, LAST ANNOTATION UPDATE)		
DE	SULIYSIN.			
GN	SLY.			
OS	STREPTOCOCCUS SUTS.			
OC	BACTERIA; FIRICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;			
OC	STREPTOCOCCUS.			
RN	[1]			
RP	SEQUENCE OF 1-71 FROM N.A.			
RC	STRAIN-PL7;			
RA	SEGRS R.P.A.M., KENTER T., DE HAAN W.A.M., JACOBS A.A.C.;			
RT	"Characterization of the gene encoding suliyisin from Streptococcus			
RT	suts, and expression in field strains."			
RL	SUBMITTED (OCT-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL: Z36907; E1334304; -			
SQ	SEQUENCE 497 AA; 54850 MW; D78ABEF CRC32;			

ALIGNMENTS

Query Match	54.9%	Score 1779;	DB 2;	Length 497;
Best Local Similarity	49.8%	Pred. No. 0.00e+00;		
Matches 233;	Conservative 110;	Mismatches 123;	Indels 2;	Gaps 2;
Db	29	SKODINOFOSITYPEPOLITNEGEYIDNPATITGMENGFVYLREKKITNNKSDIA	88	
Qy	2	AKKAVNDPILAMNDKKLTHQGESIENREXKGNOLPDEFFVXERKKRLSTNTDIX	61	
Db	89	VYDAAANITYPGALRAQNTLNDNPITLSIARGDLTSLMPELANGSDITVNSPTRS	148	
Qy	62	VATNDSRLYPALLVYDETLEKNNPILAVDRAPMTSIDLPBLASSDSITLQVEDS	121	
Db	149	TYRGTNNLISKWNTYAGEYGNTOALQYDETMAYSQSLKTFEGTSFEKIAVPLDINF	208	
Qy	122	SVRGAVNDLLAKWHDY-GQVNNVPARKQYKETAHSMELQKVFSGDFEXTGNSLIDF	180	
Db	209	DAVNGEKOYIVNKOIYTVSVDEPSPSKLFAEGTVDLRNGITDVPVYSSV	266	
Qy	181	NSVHSGEKQIQIVNNKQIIITVYSDAVANNPDGFQDTYVDELKQGISERPLVYISV	240	
Db	269	SYGRMFKLETSSRSSTVOVAAAKAIGVDISGNAEYODILKNTSAYFFGDSAGSAA	328	
Qy	241	AGRGVYIKLETTSKSEVEAAFEALIKGVAVAPOTENKQILDITVYKAVILGSDPSSGA	300	

[illegible]

RESULT	2	PRELIMINARY;	PRT;	497 AA.
ID	085102			
AC	085102;			
DT	01-NOV-1998	(TREMBL.REL. 08,	CREATED)	
DT	01-NOV-1998	(TREMBL.REL. 08,	LAST SOURCE UPDATE)	
DT	01-NOV-1998	(TREMBL.REL. 08,	LAST ANNOTATION UPDATE)	
NP	HEMOLYSIN.			
OS	STREPTOCOCCUS SUTS.			
OC	BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;			
OC	STREPTOCOCCUS.			
OC	(1)			
NP	SEQUENCE FROM N.A.			
NP	STRAIN-1933;			
RC	STRAIN-1933;			
RA	OKUMYABUA O.;			
RT	"streptococcus suis type 2 hemolysin (suiIysin) gene."			
RT	SUBMITTED (JUL-1998)	TO EMBL/GENBANK/DBD	DATA BANKS.	
NP	EMBL: AF043556; G3309561,			
SO	SEQUENCE	497 AA; 34778 MW; 36DAB184	CRC32;	

Query Match	54.7%;	Score 1774;	DB 2;	Length 497;
Best Local Similarity	49.8%;	Pred. No. 0.00e+00;		
Matches	233;	Conservative	109;	Mismatches 124;
			Indels	2;
			Gaps	2;

Db	29	SKODINOFOSLJTFGPBQIEILINEEYEDINPATQMGLENCFYILREKKNINNSADIA	88
Qy	2	ANKAVNDPIILAMNDKKKLLIHQSESIENRXXKKGNDLPDEFVYXKKKKSLSINTSDIX	61
Db	89	VIDAKKANIPGALLLRADQNLNDNPTLISITANGDLTILNPLGANGDSHTVYNSPTRS	148
Qy	62	VTANDESRLEYGALLLVYDETFEXNNPPLTLAVDRAPMTYSIDLPGLASDSFLOYEDPSNS	121
Db	149	TVRTGVNNLISKNNNTVAGEGNCITQAELODEMTAYMSOLKTKFQGSFEKIAVPLDINF	208
Qy	122	SVRAVNDLAKKHODY-GOVNANPARXQYEKITAHSMEOLKAFEGSDEKFTGNSLIDFR	180
Db	209	DAVNSGEKOVIVNFKOIIYTVSVDEDESPSKLFEAGSTVEDLKRNSITDEVPVYVSV	268
Qy	181	NSVSHSGEKOIYVXKQOIIYTVSVDAVKNPQGVFQDIYVYEDLQKORISMERPLVYISXV	240
Db	269	SYGSMFKLETSSRSRTVOOAFALAIKVDISGNAEYODILKNTSAYIFGCDAGSAA	328
Qy	241	AYGGVOYVILKETTTSKSEVEAFALIKGVKVAQOTEMKQILNDTEVKAAVILGDSGSGA	300
Db	329	TVBSGNIEETLKRIIEEGARYGKLNPGVISTSVFVKNRPAOLISXEYIEFTSTVHNS	388
Qy	301	RVTYKGDVMEDDLQOBGSRFLADHPGLPITSTTFELNDNVYATFQNSDTDYELKIVYIYRN	360
Db	389	SALTLDHSGAVAKYINTWEEVSYNEAGEEVEWPKADKNGVNLTSWSETIOIPGNARN	448
Qy	361	GDLLDHSGAVAOYIYTMNELSYDHOGKEVILAPKADMRNGQDLTAFTTISIPLKGVNRN	420
Db	449	LHVNIQETGLAMEMRTYVDK-DLPIVQGKRIINQSTIYIPQVADYV	495
Qy	421	LSVAKIRETGILAMEMKRIYVEKTDLPVRKRTISITWFTTIPQVEDYV	468

RESULT	3		
ID	031241	PRELIMINARY;	PRT; 534 AA
AC	031241;		

DT 01-JAN-1998 (TREMBLREL, 05, CREATED)
DT 01-JAN-1998 (TREMBLREL, 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL, 08, LAST ANNOTATION UPDATE)
DE PYOLYSIN.
GN PLO.
OS ARCANOBACTERIUM PYOGENES.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
NC ACTINOMYCETALES; ACTINOMYCINAE; ACTINOMYCETACEAE; ARCANOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BBR1.
RX MEDLINE; 97464437.
RA BILLINGTON S.J., JOST B.H., CUEVAS W.A., BRIGHT K.R., SONGER J.G.;
RT "The Arcanobacterium (Actinomycetes) pyogenes hemolysin, pyolysin, is a
RT novel member of the thiol-activated cytolysin family.";
RL J. BACTERIOLOG. 179:6100-6106(1997).
DR EMBL; U84782; G252800; -;
DR PFAM; PF01289; Thiol_cytolysin; 1.
SQ SEQUENCE 534 AA; 57873 MW; B0C14CA8 CRC32;

Query Match	40.2%;	Score 1302;	DB 2;	Length 534;
Best Local Similarity	40.0%;	Pred. No. 1.55e-248;		
Matches	189;	Conservative 121;	Mismatches 154;	Indels 8;
				Gaps 6;

[illegible]

RESULT	ID	4	PRELIMINARY;	PRT;	50 AA.
AC	Q48772.				
DT	01-NOV-1996	(TREMBLEL, 01, CREATED)			
DT	01-NOV-1996	(TREMBLEL, 01, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998	(TREMBLEL, 08, LAST ANNOTATION UPDATE)			
DE	LISTERIOLYSIN O (FRAGMENT).				
GN	HELIX.				
OS	LISTERIA MONOCYTOGENES.				
OC	BACTERIA; FIRMICUTES, BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE.				
CC	LISTERIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=11984, TYPE 1;				
FX	MEDLINE; 96118685.				
RA	RAMMOSEN O.F., SKOUBOE P., DONS L., ROSSEN L., OLSEN J.E.;				

SO	SEQUENCE	748	AA;	84361	MM;	4900A1F6	CRC32;
Query Match		3.5%	Score 112;	DB 3;	Length 748		
Best Local Alignment		10.8%	Score 127	DB 3	Length 748		

```
Query Match      3.5%; Score 112; DB 3; Length 748;
Best Local Similarity 18.8%; Pred. No. 8.43e-02;
```

Matches	33;	Conservative	58;	Mismatches	71;	Indels	14;	Gaps	13;	
Db	567	SEEEETEDDDPFCIADIOLVYSSKISTSDSQNNTTSDDIIDTSAASSIASPEFCIIM	626							
Qy	22	THGGS-IEHFxxEGNLPDEFVxxERKRSLSINTSDIXVTATNSRLYPGAL--Lvv	78							
Db	627	SOSMKEILQSJKTVGLKMRKTRKEVILQSOALSTANPDNGEGGVANF-SK-IEI	684							
Qy	79	DETLKENPPTLLADRAPMTYSIDL-PELASSDSFLOYEDPSNSVKGAVNDLAKKHOD	137							
Db	665	FDHLELLIEAEPDFLEIRTYTEPIPLNELIKLSAEPFVSQIDEMTIREWADVQG	740							
Qy	138	YGVVNN-VPARQY-EKI-T-AH-SMEQLKVK-EGSD-FEKTGNSLDI-DENSVHS	185							
RESULT	8									
ID	087083	PRELIMINARY;	PRT;	1361	AA.					
AC	087083;									
DT	01-NOV-1998	(TREMBLREL. 08, CREATED)								
DT	01-NOV-1998	(TREMBLREL. 08, LAST SEQUENCE UPDATE)								
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)								
DE	S-LAYER PROTEIN.									
OS	CAMPYLOBACTER RECTUS.									
OC	BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;									
CC	CAMPYLOBACTER.									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN-ATCC 33238;									
RX	MEDLINE; 98442421.									
RA	MUYAMOTO M., MAEDA H., KITANAKA M., KOKEGUCHI S., TAKASHIBA S.,									
RA	MUYAYAMA Y.;									
RT	"The S-layer protein from Campylobacter rectus: sequence									
RT	determination and function of the recombinant protein.";									
RL	FEBS MICROBIOL. LETT. 166:275-281(1998).									
DR	EMBL; AB001876; D1034498; -.									
KW	S-LAYER.									
SO	SEQUENCE	1361	AA;	144904	MM;	55A9D736	CRC32;			
	Query Match	3.5%;	Score	113;	DB	2;	Length	1361;		
	Best Local Similarity	21.0%;	Pred.	No.	6.18e-02;					
	Matches	39;	Conservative	49;	Mismatches	88;	Indels	10;	Gaps	10;
Db	178	VDWHSNPEHPKAYELTTNTDNTATNANPNMK-HNPGGTORITLQSDKLTJGDYSRD	236							
Qy	115	VEDPSSNVSRGAVNDLAKMHODYGVQVNNVPARQYERKITAHSM-EQLKVFEGSDFEKYG	173							
Db	237	NLNAEFQOANADEGDPRTPTLTINIONINIEVYGTNTADL-RDSNDVAKINIHRTK	295							
Qy	174	NSLDIDFNSVHSGE-KQI-QIVNKQI-YIYVSVDAYANPGDYQDYTYVEDLKQRGISA	230							
Db	296	EAGNKFENYESIG-OKLVGMRLANVAK-KDIDYKEH-KKGLSGFEDKSNVLENEAKS	352							
Qy	231	ERPLYV-ISXVAYGROYLKLLETTSKSXEVEAFAELIKGVAPQTEWKOILNTEVKA	289							
Db	353	LSITSD	358							
Qy	290	VILGSD	295							
RESULT	9									
ID	030524	PRELIMINARY;	PRT;	1361	AA.					
AC	030524;									
DT	01-JAN-1998	(TREMBLREL. 05, CREATED)								
DT	01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)								
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)								
DE	S-LAYER PROTEIN.									
GN	CRS.									
OS	CAMPYLOBACTER RECTUS.									
OC	BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;									
CC	CAMPYLOBACTER.									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN-314;									

```

RA MEDLINE: 98187925.
RX WANG B., KRAIG E., KOLODRUBETZ D.;
RT "A new member of the S-layer protein family: characterization of the
RT cis gene from Campylobacter rectus.";
RL INECT. IMMUN. 66:1521-1526(1998).
DR EMBL: AF010143: G2459961; ".
SQ SEQUENCE 1361 AA; 144385 MW; CAPE081F CRC32;

Query Match 3.5%; Score 113; DB 2; Length 1361;
Best Local Similarity 21.0%; Pred. No. 6,18e-02;
Matches 39; Conservative 49; Mismatches 88; Indels 10; Gaps 10;

Db 178 VDMNSPEHPKAVELTNTDNTANAFENAPMK-HNPGGDRIMTLQSDKLTGDSYRHD 236
OY 115 VEDPSNSSVRCAVNDLAKKHQDYGYNNVPARKQYEKIPAHSMF-OLKKYFGSDFEKTG 173
Db 237 NTLNVEFGQANADEGDPSTRPTLTNTQININIEVTGVTNLTDL-RDSNDYEKINIRITK 295
OY 174 NSLIDIDFNSVHSGE-KQI-QIVNK-KQI-YIVVSVDYAKNPEDVQDPIVYVTELDKQRGISA 230
Db 296 EAGKFNVEISG-OKYGMRLANNAK-KDIDYKEH-KGVLSGFEDKSNVLELVNAKS 352
OY 231 ERPLTY-ISKVAVGYLYLKLETTSKSXEVAEALIKGVKAPQREMKQILDNTEVKA 289
Db 353 LSTSD 358
OY 290 VILGD 295

RESULT 10 PRELIMINARY; PRT; 187 AA.
AC 058072.
DT 01-AUG-1998 (TREMBLREL, 07, CREATED)
DI 01-AUG-1998 (TREMBLREL, 07, LAST SEQUENCE UPDATE)
DE 01-JAN-1999 (TREMBLREL, 09, LAST ANNOTATION UPDATE)
DE 187AA LONG HYPOTHETICAL PROTEIN.
CN PH0334.
OS PYROCOCCLUS HORIKOSHII.
OC ARCHAEA; EURARCHAEOTA; THERMOCOCCALES; THERMOCOCCACEAE; PYROCOCCLUS.
RN 11]
RP SEQUENCE FROM N.A.
RC STRAIN-OT3;
RX MEDLINE: 98344137.
RA KAWARABAYAST Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOTAMA A., NAGAI Y.,
RA SAKAI M., OGURA K., OJUKA R., NAKAZAWA H., TAKAMURA M., OHFUKU Y.,
RA FUKAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,
RA AOKI K., NAKAMURA Y., ROBE T.F., HORIKOSHI K., MASUOCHI Y., SHIZUYA H.,
RA KIRUCHI H.;
RT "Complete Sequence and Gene Organization of the Genome of a
RT Hyper-thermophilic Archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA RES. 5:55-76(1998).
DR EMBL: AP000002; D1030351; ".
SQ SEQUENCE 187 AA; 21959 MW; BEC990E3 CRC32;

Query Match 3.4%; Score 111; DB 1; Length 187;
Best Local Similarity 32.2%; Pred. No. 1,15e-01;
Matches 19; Conservative 13; Mismatches 21; Indels 6; Gaps 6;

Db 117 LLNGAFIYEQO-YGLTKVI-K-K-VDNGYGLALITEKKIIPYVDWKVY-ENPIKIRV 170
OY 234 LVYISVAVGYLYLKLETTSKSXEVAEAF-ALIKGVKAPQREMKQILDNTEVKA 291

RESULT 11 PRELIMINARY; PRT; 814 AA.
AC 065715.
DT 01-AUG-1998 (TREMBLREL, 07, CREATED)
DI 01-AUG-1998 (TREMBLREL, 07, LAST SEQUENCE UPDATE)
DE 01-JAN-1999 (TREMBLREL, 09, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 91.9 KD PROTEIN.
TX 5TK18.220.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).

```

OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
 OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
 OC CAPRIFALES; BRASSICACEAE; ARABIDOPSIS.

RA BEVAN M., VAN DER SCHUEREN J., CHUANG Y.-J., VOET M., ROBBEN J.,
 RA VOLCKERT G., BANCROFT I., MEWES H.W., MAYER K.F.X., SCHUELLER C.,
 RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]

RA EU ARABIDOPSIS SEQUENCING PROJECT;
 RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; AL022580; E1287635; -
 KW HYPOTHETICAL PROTEIN.

SO SEQUENCE 814 AA; 91943 MW; A418987 CRC32;
 Query Match
 Best Local Similarity 3.4%; Score 111; DB 10; Length 814;
 Matches 20; Conservative 29; Mismatches 30; Indels 5; Gaps 5;

Db 729 EMBKNNVHPIKITYVMIGYARQNTGASRLNEMREKIVPD-SITKEFIYGLKQ 787
 QY 187 EKQIQVKKQIYTVSDANKNGDFQDTVYEDLKQGISARPLYVTSXV-AYGRQ 245
 Db 788 GGV-LEA-FKGS-D-EENYALIEG 808
 QY 245 VYLKLETTSKSXVEAFAEALIKG 269

RESULT 12
 ID 058691 PRELIMINARY; PRT; 218 AA.
 AC 058691;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL PROTEIN M01295.
 GN M01295.
 OS METHANOCOCCUS JANNASCHII.
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
 OC METHANOCOCCUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE; 96337999.
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
 RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,
 RA SCOTT J.L., GEORGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
 RA COTTON M.D., ROBERTS C.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.,
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii";
 RL SCIENCE 273:1058-1073(1996).
 CC -1- SIMILARITY: STRONG, TO A.FULGIDUS AF1550 AND
 CC M.THERMOAUTOTROPHICUM MTH576.
 DR EMBL; 067570; G1591933; -
 KW HYPOTHETICAL PROTEIN.

SO SEQUENCE 218 AA; 24666 MW; CFC99A99 CRC32;
 Query Match
 Best Local Similarity 3.3%; Score 107; DB 1; Length 218;
 Matches 21; Conservative 17; Mismatches 40; Indels 2; Gaps 2;

Db 68 IGVGLTHIPKPSIPKAKKADLGAIEIVYH-GEIVVEVEKTYTYSISDDVILAH 126
 QY 153 ITHASHQQLVKKGSDEKFGNSLDIDFNSVHSGEKQIQIVNKKQIYTVSDAVKPNPD 212
 Db 127 PGFIDETAKENKENDIVE 146
 QY 213 V-FQDTVYEDLKQGISAE 231

RESULT 13
 ID 012044 PRELIMINARY; PRT; 265 AA.
 AC 012044;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
 DE CHROMOSOME XV READING FRAME ORF YOR220W.
 OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCETES; SACHAROMYCETACEAE;
 OC SACHAROMYCETACEAE; SACHAROMYCES.

RA BOYER J., FAIRHEAD C., GAILLON L., GALLISSON F., MICHAUX G.,
 RA THIERRY A., DUJON B.,
 RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA BOYER J., FAIRHEAD C., GAILLON L., GALLISSON F., MICHAUX G.,
 RA THIERRY A., DUJON B.,
 RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FY1679;
 RA GALLISSON F., DUJON B.,
 RL SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; 275128; E252085; -
 DR EMBL; X92441; G1050772; -
 SO SEQUENCE 265 AA; 29255 MW; 2D5DD78E CRC32;

Query Match
 Best Local Similarity 3.3%; Score 106; DB 3; Length 265;
 Matches 26; Conservative 26; Mismatches 44; Indels 7; Gaps 7;

Db 85 PKKLEVTSLNNQ-HPR-SRSTDAVSLQDNNLALDEHKNKPLSTNTPGYTGDDSS 142
 QY 12 ANN-YDKKKLTHQGESEIENKXKNQLPDE-IVYAE-RKKRSISTSDIXVATINDS 68
 Db 143 SLNKGSSLSPPKSSLESPYMKLSTDSKPSYQEPPLKLSRSS 187
 QY 69 RLY-EGALLVDETLKXNNPTL-LAVDRAPMTYSIDPLGLASSDS 111

RESULT 14
 ID 028947 PRELIMINARY; PRT; 501 AA.
 AC 028947;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN AF1322
 OS ARCHAEoglobus fulgidus.
 OC ARCHAEA; EURYARCHAEOTA; ARCHAEoglobales; ARCHAEoglobaceae;
 OC ARCHAEoglobus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VG-16 / DSM 4304 / ATCC 49558;
 RA MEDLINE; 98049343.
 RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,
 RA KETCHUM K.A., DODSON R.J., GWINN M., HICKY E.K., PETERSON J.D.,
 RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KIRKNESS N.C.,
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
 RA PETERSON S., REICH C.I., MCNEIL L.R., BADGER J.H., GLODER A., ZHOU L.,
 RA OVERBEK R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,
 RA COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYRES S.M.,
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,
 RA VENTER J.C.,
 RT "The complete genome sequence of the hyperthermophilic,
 RT sulphate-reducing archaeon Archaeoglobus fulgidus";
 RL NATURE 390:364-370(1997).
 DR EMBL; AE001012; G2649254; -
 DR TIGR; AF1322; -
 KW HYPOTHETICAL PROTEIN.

SQ SEQUENCE 501 AA; 55698 MW; 5EE0081E CRC32;

Query Match 3.3%; Score 106; DB 1; Length 501;

Best Local Similarity 27.5%; Pred. No. 5.25e-01;

Matches 22; Conservative 24; Mismatches 31; Indels 3; Gaps 3;

Db 217 LVAVNDSGVGRIVE-VNGYKVSVDVAKAVLONAEIAEIKIYNGDEIRILSYPAVMGV 275

QY 250 LETTSKSEVEAEAFALIKGVKAPOTEMKOILDNTEVKAV-ILGDPSSGARVYTGKVD 308

Db 276 KVIGLYTENGEKPEAPLAGI 295

QY 309 WVEDL-IOEGSRFTADHPGL 327

RESULT 15
ID 085906 PRELIMINARY; PRT; 710 AA.

AC 085906; (TREMBLEREL. 08, CREATED)

DT 01-NOV-1998 (TREMBLEREL. 08, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMBLEREL. 08, LAST ANNOTATION UPDATE)

DE TRANSPONSON PROTEIN B.

GN TNBP.

OS SPHINGOMONAS AROMATICIVORANS.

OG PLASMID PVL1.

OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; ZYMOMONAS GROUP;

OC SPHINGOMONAS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-F199;

RA ROMINE M.F., STILLWELL L.C., MONG K.-K., THURSTON S.J., SISK E.C.,

RA SENSEN C.W., GAASTERLAND T., SAFFER J.D., FREDRICKSON J.K.,

RT "Complete sequence of a 184 kb catapolic plasmid from SpHingomonas

aromaticivorans strain F199."

RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

RE EMBL: AF079317; G3378338; -

RE PLASMID.

RE SEQUENCE 710 AA; 81079 MW; 41B48670 CRC32;

SQ SEQUENCE 710 AA; 81079 MW; 41B48670 CRC32;

Query Match 3.3%; Score 107; DB 2; Length 710;

Best Local Similarity 22.6%; Pred. No. 3.89e-01;

Matches 14; Conservative 20; Mismatches 25; Indels 3; Gaps 3;

Db 139 NHHNDVLECYAMYGANGRR-RQSQT-RQMRGKIVTSEFRSPYQLFGDIHRAII 196

QY 365 DHSGAYVAQYITWNLSEYDHOGKEVLTPKAWDRNGODLTAHFTTS-IPKGNVRLISVK 424

Db 197 NR 198

QY 425 IR 426

Search completed: Mon Aug 30 16:01:04 1999

Job time : 79 secs.


```

Db 181 nsyhsgekgdqvnfkglyyrvsvdavnpgdvfgdvtvedlkgqrsaeprlvysv 240
Oy 181 NSVHSGEKGQIOIVNFKQIYTVSVDAVNPGDVFGDVTVEDLKQRGISAEPRLVYISSV 240
Db 241 aysrgvylkletsksdeveaafaalngvavapotekqjldntevkavllggdpsga 300
Oy 241 AYSGRVYVLEKLETTSSKDSDEVAEAFALNGVAVAPOTEKQJLDNTEVKAVLLGGDPSGA 300
Db 301 rvtvgkvmdvedllqesgfadhpjlpisyltsfjrdnvaatfgnsdtyvetkvtayrn 360
Oy 301 RVVTGKVMVEDLLQESGFRFADHPGLPISTYTSFLRDNVAVATQNSDYEKVTAYARN 360
Db 361 gdlldhsagaayaqyitwlnelSYDHQKEVLTFRAMRNOGDLFAHFTTSIPKGNVRN 420
Oy 361 GDLLDHSAGAAYAQYIYTWLNELSYDHQKEVLTFRAMRNOGDLFAHFTTSIPKGNVRN 420
Db 421 lsvkirectglawewrtvyekeidplvrkrttsimgtllpyvedkvend 471
Oy 421 LSVKIRECTGLAWEMWRTVYEKIDPLVRKRTISIMGTLLPYVEDKVENND 471

```

```

RESULT 2
ID R05924 standard; protein; 471 AA.
AC R05924;
DE 23-NOV-1990 (first entry)
KW Immunogenic pneumolysin variant.
KW pneumolysin; vaccine; pneumonia; meningitis; bacteriaemia; ds.
OS Streptococcus pneumoniae.
FH Key Location/Qualifiers
FT misc-difference 428
FT /label=G, A, S
FT /label=Q, D
FT W09006951-A.
FN 28-JUN-1990.
PD 15-DEC-1989; AU0539.
PR 16-DEC-1988; AU-001989.
PA (PATO/) PATON J C.
PI PATON JC, HANSMAN DJ, BOULNOIS GJ, ANDREW PW, MITCHELL TJ,
PI WALKER JA;
DR WPI; 90-224494/29.
PT New non-toxic, immunogenic mutants of pneumolysin - useful in
PT protective vaccines against Streptococcus pneumoniae, and DNA
PT sequences encoding them.
PS Disclosures, P; English.
CC Vaccines are non-toxic and antigenic to wild type pneumolysin,
CC making them useful in vaccination against pneumonia and associated
CC meningitis, bacteriaemia etc.
SQ Sequence 471 AA;

```

```

Query Match 94.6%; Score 3121; DB 1; Length 471;
Best Local Similarity 96.4%; Pred. No. 2.05e-259;
Matches 454; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Db 1 manxavdflamnydkkllthbgsienrfikgnqldpdefvlerkrkststsd 60
Oy 1 MANXAVDFILAMNYDKKLLTHBGSIENRFIKGNQLPDEFVLERKRKSLSTSTSDI 60
Db 61 svratndsrlypgallvvdetllennptllavdrapmtysidlpjlassdsfivgedpsn 120
Oy 61 SVRATNDSRLYPGALLVVDETLLENNPTLLAVDRAPMTYSIDLPJGLASSDSFQVGEDPSN 120
Db 121 ssrvgaavndllakwgdqgygvnvarmyekktahamegkjkvfgsdfektynslidf 180
Oy 121 SSVRGAAVNDLLAKWHDQYGVNVPARMQYKKTASHMEQKVKVFGSDEKTSNDIDF 180
Db 181 nsyhsgekgdqvnfkglyyrvsvdavnpgdvfgdvtvedlkgqrsaeprlvysv 240
Oy 181 NSVHSGEKGQIOIVNFKQIYTVSVDAVNPGDVFGDVTVEDLKQRGISAEPRLVYISSV 240
Db 241 aysrgvylkletsksdeveaafaalngvavapotekqjldntevkavllggdpsga 300
Oy 241 AYSGRVYVLEKLETTSSKDSDEVAEAFALNGVAVAPOTEKQJLDNTEVKAVLLGGDPSGA 300

```

```

Oy 241 AYSGRVYVLEKLETTSSKDSDEVAEAFALNGVAVAPOTEKQJLDNTEVKAVLLGGDPSGA 300
Db 301 rvtvgkvmdvedllqesgfadhpjlpisyltsfjrdnvaatfgnsdtyvetkvtayrn 360
Oy 301 RVVTGKVMVEDLLQESGFRFADHPGLPISTYTSFLRDNVAVATQNSDYEKVTAYARN 360
Db 361 gdlldhsagaayaqyitwlnelSYDHQKEVLTFRAMRNOGDLFAHFTTSIPKGNVRN 420
Oy 361 GDLLDHSAGAAYAQYIYTWLNELSYDHQKEVLTFRAMRNOGDLFAHFTTSIPKGNVRN 420
Db 421 lsvkirectglawefrtvyekeidplvrkrttsimgtllpyvedkvend 471
Oy 421 LSVKIRECTGLAWEMWRTVYEKIDPLVRKRTISIMGTLLPYVEDKVENND 471

```

```

RESULT 3
ID R33842 standard; protein; 480 AA.
AC R33842;
DE 15-JUL-1993 (first entry)
KW Streptolysin O variant mSLO.3/6.
KW SLO; soluble; haemolytic activity; wild type; anti-SLO antibodies;
KW Streptococcus pyogenes; ASO.
OS Streptococcus pyogenes.
FH W09305155-A.
FN 18-MAR-1993;
PD 03-AUG-1992; U06380.
PR 30-AUG-1991; US-752428.
PA (BECI ) BECKMAN INSTR INC.
PI Adams CW;
DR WPI; 93-100979/12.
DR N-FSD8; Q38287.
PT Streptolysin O variants produced by recombinant DNA technology -
PT having no haemolytic activity and recognised by wild-type
PT anti-streptolysin O antibodies, useful in diagnosis of
PT Streptococcus pyogenes infection
PS Claim 17; Fig 2; 80pp; English.
CC The sequence represents a soluble variant of Streptolysin O (SLO)
CC having no haemolytic activity. The variant SLO can specifically
CC recognise wild-type anti-SLO antibodies (ASO) but has no haemolytic
CC activity. It can be used for the determination of previous and/or
CC current infection by Streptococcus pyogenes. It enables ASO assays
CC to become commercially viable and avoids the risk of handling haemo-
CC lytically active S. pyogenes wild-type SLO.
SQ Sequence 480 AA;

```

```

Query Match 43.8%; Score 1444; DB 6; Length 480;
Best Local Similarity 42.8%; Pred. No. 7.18e-112;
Matches 196; Conservative 108; Mismatches 151; Indels 3; Gaps 3;

Db 18 lndkylslnyelevlakngelien-fvpekyvkkadkfivlerkknkntcpvdslid 76
Oy 18 LNDKYLSPGALLVVDETLLENNPTLLAVDRAPMTYSIDLPJGLASSDSFQVGEDPSN 76
Db 65 tndsrlypgallvvdetllennptllavdrapmtysidlpjlassdsfivgedpsn 124
Oy 65 TNDSRLYPGALLVVDETLLENNPTLLAVDRAPMTYSIDLPJGLASSDSFQVGEDPSN 124
Db 136 taldnlvqwhdnygsncltpartqylksmynyskqleaalnvnsklldtldgfsis 195
Oy 136 GAVNDLLAKWHDQYGVNVPARMQYKKTASHMEQKVKVFGSDEKTSNDIDFNSVH 184
Db 196 kgekvmiaaykqifytvsanlpnpadvdkavtkelrkqsvsneapplfvsnaygr 255
Oy 196 KGEKVMIAAYKQIFYTVSANLPNPADVDKAVTKELRKQSVSNEAPPLFVSNAYGR 255
Db 256 tvyfkleksskndveaafaalngvavapotekqjldntevkavllggdpsga 315
Oy 256 TVYFKLEKSSKNDVEAFAALNGVAVAPOTEKQJLDNTEVKAVLLGGDPSGARVYT 304
Db 316 kdfavlnvkladnatfcrnpayisyltsfjrdnvaatfgnsdtyvetkvtayrn 375
Oy 316 KDFAVLNVLKADNATFCRNPAYISYLSFJRDNVAATQNSDYEKVTAYARN 375
Db 305 gkvmdvedllqesgfadhpjlpisyltsfjrdnvaatfgnsdtyvetkvtayrn 420
Oy 305 GKVMVEDLLQESGFRFADHPGLPISTYTSFLRDNVAVATQNSDYEKVTAYARN 420

```


[illegible][illegible]

Db 200 ydltv-amkar-ev-qdaykaenldkakaavdinglypkvdaftelt-evakkaida 255
 QY 200 YTVSDAVKNNPGVFDVTDVDELKORGISAERPLVYISSVAYGRQVYLKLETTSKS-DE 258
 Db 256 deaalpkrvesv-aatgpkaveltavpugt-lkglisaanedvnyvnyrykxdg 313
 QY 259 VEAFAEALFKGVAVAGQTEMKQ-LDNTVEKAVILGDDSSGARVYTGAKDVEDLIDEG 317
 Db 314 nlpfalntadsvlsldgkltvdaasptfienkykvkqkdkng 359
 QY 318 S-RFTADHPLPISYTTSTFLRDNVATFONSTDY-VEIKVATYRNG 361

RESULT 12

ID W55347 standard; Protein: 188 AA.
 AC W55347;
 DT 17-JUN-1998 (first entry)
 DE H. pylori ORF 01cel1618orf1 protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KM Identification; binding compound; bacteria; life cycle; activator;
 OS Inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 PN Helicobacter pylori.
 PD W09737044-A1.
 PF 09-OCT-1997.
 PR 27-MAR-1997; U05223.
 PR 06-DEC-1996; US-761318.
 PR 29-MAR-1996; US-625811.
 PR 02-APR-1996; US-758731.
 PR 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D.
 DR WPI: 97-503122/46.
 N-PSDB: V24736.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PS infection and for diagnosis of H. pylori infection
 PS Claim 14; Page 571; 1145pp; English.
 CC This sequence is a H. pylori protein of unspecified function.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SQ Sequence 188 AA;

Query Match 3.2%; Score 104; DB 29; Length 188;
 Best Local Similarity 34.4%; Pred. No. 2.32e+01;
 Matches 21; Conservative 21; Mismatches 16; Indels 3; Gaps 3;
 Db 95 psslmrdsvsldskkr-lnaikdlifhknkf-rqlqlkintpika-lveagkdgfka 151
 QY 210 PGDVPDVTYVEDLKGKRGISAERPLVYISSVAYGRQVYLKLETTSKSDEVAEFAELIKG 269
 Db 152 1 152
 QY 270 V 270

RESULT 13
 ID W5528 standard; Protein: 418 AA.
 AC W5528;
 DT 02-JUL-1998 (first entry)
 DE H. pylori ORF 299630321_24336712_f1_5 cellular protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KM Identification; binding compound; bacteria; life cycle; activator;
 OS Inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 PN Helicobacter pylori.
 PD W09737044-A1.
 PF 09-OCT-1997.
 PR 27-MAR-1997; U05223.
 PR 06-DEC-1996; US-761318.
 PR 29-MAR-1996; US-625811.
 PR 02-APR-1996; US-758731.
 PR 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D.
 DR WPI: 97-503122/46.
 N-PSDB: V24937.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PS infection and for diagnosis of H. pylori infection
 PS Claims 14,93; Page 730-731; 1145pp; English.
 CC This sequence is a H. pylori cellular protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SQ Sequence 418 AA;

Query Match 3.2%; Score 104; DB 29; Length 418;
 Best Local Similarity 34.4%; Pred. No. 2.32e+01;
 Matches 21; Conservative 21; Mismatches 16; Indels 3; Gaps 3;
 Db 325 psslmrdsvsldskkr-lnaikdlifhknkf-rqlqlkintpika-lveagkdgfka 381
 QY 210 PGDVPDVTYVEDLKGKRGISAERPLVYISSVAYGRQVYLKLETTSKSDEVAEFAELIKG 269
 Db 382 1 382
 QY 270 V 270

RESULT 14
 ID R80623 standard; Protein: 210 AA.
 AC R80623;
 DT 12-MAR-1996 (first entry)
 DE Coho salmon growth hormone.
 KW Growth hormone; coho salmon growth hormone; cSH; transgenic; abalone;
 KM actin; ACT5-C; insulin-like growth factor; cytochalasin B;
 OS Oncorhynchus keta.
 FH Key
 FT peptide
 FT 1..22 Location/Qualifiers
 FT /note= "signal peptide"
 FT 23..210

Db 121 SSVGAANDLLAKWHDYGOVNNVPAARMQYEKITAHSMEOLKVKFGSDPEKTSNLSIDIDF 180
 |||||
 Qy 121 SSVGAANDLLAKWHDYGOVNNVPAARMQYEKITAHSMEOLKVKFGSDPEKTSNLSIDIDF 180
 Db 181 NSVHSGEKOIOIVNFKOIYTVSVDAVKNPGDYFODIVTYEDLKORISAEPLVYISSV 240
 |||||
 Qy 181 NSVHSGEKOIOIVNFKOIYTVSVDAVKNPGDYFODIVTYEDLKORISAEPLVYISSV 240
 Db 241 AYGRQVYLKLETTSSKSDVEAFAEALIKGVKVAPOTEWKOILNTEVKAVIIGDPSGA 300
 |||||
 Qy 241 AYGRQVYLKLETTSSKSDVEAFAEALIKGVKVAPOTEWKOILNTEVKAVIIGDPSGA 300
 Db 301 RVYTGKVMEDDIOEGSRFTADHPGLPISTYSFLRDNVNVAEONSTDVETKVAAYRN 360
 |||||
 Qy 301 RVYTGKVMEDDIOEGSRFTADHPGLPISTYSFLRDNVNVAEONSTDVETKVAAYRN 360
 Db 361 GDLILDHSGAVYAOYIYTWDELSDYHOGKEVLTLPKAMDNRGODLTAHFTTSIDPKGNVRN 420
 |||||
 Qy 361 GDLILDHSGAVYAOYIYTWDELSDYHOGKEVLTLPKAMDNRGODLTAHFTTSIDPKGNVRN 420
 Db 421 LSVKIRECTGLAMEMWRTVEKTDPLVRRKRTISIGTTLYPQVEDKVEN 471
 |||||
 Qy 421 LSVKIRECTGLAMEMWRTVEKTDPLVRRKRTISIGTTLYPQVEDKVEN 471

RESULT 2
 ENTRY B43577 #type complete
 TITLE perfringolysin O precursor - Clostridium perfringens
 ORGANISM #formal_name Clostridium perfringens
 DATE 03-Feb-1993 #sequence_revision 03-Mar-1993 #text_change 13-Sep-1998
 ACCESSIONS B43577; A34951; A60922

REFERENCE
 #authors Shimizu, T.; Okabe, A.; Minami, J.; Hayashi, H.
 #journal Infect. Immun. (1991) 59:137-142
 #title An upstream regulatory sequence stimulates expression of the perfringolysin O gene of Clostridium perfringens.
 #cross-references M3577; 91099951

REFERENCE
 #authors Tweten, R.K.
 #journal Infect. Immun. (1988) 56:3235-3240
 #title Nucleotide sequence of the gene for perfringolysin O (the beta-toxin) from Clostridium perfringens: significant homology with the genes for streptolysin O and pneumolysin.
 #cross-references M3577; 89032623

REFERENCE
 #authors Tweten, R.K.
 #journal Infect. Immun. (1988) 56:3228-3234
 #title Cloning and expression in Escherichia coli of the perfringolysin O (beta-toxin) gene from Clostridium perfringens and characterization of the gene product.
 #cross-references M3577; 89032622

REFERENCE
 #authors Tweten, R.K.
 #journal Infect. Immun. (1988) 56:3228-3234
 #title Cloning and expression in Escherichia coli of the perfringolysin O (beta-toxin) gene from Clostridium perfringens and characterization of the gene product.
 #cross-references M3577; 89032622

GENETICS
 #gene pfo; pfoA
 FEATURE
 #residues 1-28
 #product perfringolysin O status predicted #label SIG
 #experimental_source strain EGD
 #note this sequence is derived from a weakly hemolytic strain, serotype 1/2a

SUMMARY #length 500 #molecular-weight 55799 #checksum 4821

Query Match 49.0%; Score 1618; DB 2; Length 500;
 Best Local Similarity 46.9%; Pred. No. 1,096-272;
 Matches 216; Conservative 110; Mismatches 132; Indels 3; Gaps 3;

Db 35 NOSIDSGISLSYRNENVLANSNGKITS-FVPEKGRKAGKFLYVEKOKSLTSPDIS 93
 |||||
 Qy 3 KNAVDFLLANNYKKLLTHOGESIEENRFI-KEGNOLPEFVIEKRSLSSTNSDIS 61
 |||||
 Db 94 IIDSYNRTYAGALQADKAFENRPTILMKRKPININIDPLGLGENS-IRVDDPTYG 152
 |||||
 Qy 62 VTAINDSLRYGALLVYDETLLENPILLAVADAPMTYSIDPLGLASSDSFLOYEDPNS 121
 |||||
 Db 153 KVSAGIDELYSKNEKYSSTHTLPARTQYSESNVSKQISALNNAYKLENSLGVEN 212
 |||||
 Qy 122 SVRGAANDLLAKWHDYGOVNNVPAARMQYEKITAHSMEOLKVKFGSDPEKTSNLSIDIDF 181
 |||||
 Db 213 AVANNEKVMYLLAKWHDYGOVNNVPAARMQYEKITAHSMEOLKVKFGSDPEKTSNLSIDIDF 181
 |||||
 Qy 182 SVHSGEKOIOIVNFKOIYTVSVDAVKNPGDYFODIVTYEDLKORISAEPLVYISSV 241
 |||||
 Db 273 YGRITVYKLETTSSSKDVQAFAKLIKNTDIKNSQYKDIYENSFTAVVIGDQAEHNR 332
 |||||
 Qy 242 YGRQVYLKLETTSSKSDVEAFAEALIKGVKVAPOTEWKOILNTEVKAVIIGDPSGA 301
 |||||
 Db 333 VYTKDFEIRKVIKDNATSTKNAIPYSTSVFLKDNVAAVHNKTDYIETTSTYSGK 392
 |||||
 Qy 302 VYTKDFEIRKVIKDNATSTKNAIPYSTSVFLKDNVAAVHNKTDYIETTSTYSGK 361
 |||||
 Db 393 KINLDHSGAVYAOYIYTWDELSDYHOGKEVLTLPKAMDNRGODLTAHFTTSIDPKGNVRN 452
 |||||
 Qy 362 DILDHSGAVYAOYIYTWDELSDYHOGKEVLTLPKAMDNRGODLTAHFTTSIDPKGNVRN 421
 |||||

RESULT 3
 ENTRY A43505 #type complete
 TITLE listeriolysin O precursor - Listeria monocytogenes
 ORGANISM #formal_name Listeria monocytogenes
 DATE 21-Oct-1982 #sequence_revision 21-Oct-1992 #text_change 17-Mar-1999
 ACCESSIONS A43505; S05306; A47606; S12400; A61079
 REFERENCE
 #authors Mengaud, J.; Vicente, M.F.; Chenevert, J.; Pereira, J.M.; Geoffroy, C.; Gicquel-Sanzey, B.; Baquero, F.; Perez-Diaz, J.C.; Cosset, P.
 #journal Infect. Immun. (1988) 56:766-772
 #title Expression in Escherichia coli and sequence analysis of the listeriolysin O determinant of Listeria monocytogenes.
 #cross-references M3577; 88153053

REFERENCE
 #authors Domann, E.; Charraborty, T.
 #journal Nucleic Acids Res. (1989) 17:6406
 #title Nucleotide sequence of the listeriolysin gene from a Listeria monocytogenes serotype 1/2a strain.
 #cross-references M3577; 89366684

REFERENCE
 #authors Domann, E.; Charraborty, T.
 #journal Nucleic Acids Res. (1989) 17:6406
 #title Nucleotide sequence of the listeriolysin gene from a Listeria monocytogenes serotype 1/2a strain.
 #cross-references M3577; 89366684

REFERENCE
 #authors Domann, E.; Charraborty, T.
 #journal Nucleic Acids Res. (1989) 17:6406
 #title Nucleotide sequence of the listeriolysin gene from a Listeria monocytogenes serotype 1/2a strain.
 #cross-references M3577; 89366684

REFERENCE
 #authors Domann, E.; Charraborty, T.
 #journal Nucleic Acids Res. (1989) 17:6406
 #title Nucleotide sequence of the listeriolysin gene from a Listeria monocytogenes serotype 1/2a strain.
 #cross-references M3577; 89366684

ACCESSIONS	S22341; S36683
REFERENCE	S22340
#authors	Haas, A.; Dumbsky, M.; Kreft, J.
#journal	Biochim. Biophys. Acta (1992) 1130:81-84
#title	Listeriolysin genes: complete sequence of <i>llo</i> from <i>Listeria</i>
##cross-references	Ivanovii and of <i>lso</i> from <i>Listeria seeligeri</i> .
#####	MUID:92182018
#accession	S22341
##molecule_type	DNA
##residues	1-528 ##label HAS
##cross-references	EMBL:X60461
##note	the authors translated the codon ACA for residue 331 as Val
REFERENCE	S36683
#authors	Kreft, J.
#submission	submitted to the EMBL Data Library, July 1991
#accession	S36683
##molecule_type	DNA
##residues	1-319,'T',321-528 ##label KRE
##cross-references	EMBL:X60461
GENETICS	
#gene	<i>llo</i>
CLASSIFICATION	#superfamily dipeptide transport protein
FEATURE	
1-24	
25-528	
SUMMARY	
Query Match	46.7%; Score 1541; DB 2; Length 528;
Best Local Similarity	43.3%; Pred.No.5,70e-258;
Matches 207; Conservative 121; Mismatches 142; Indels 1; Gaps 1,	
Db	60 IDQYIGLDYDKNNILVVDGSAVKWVPKAKYKGGNOYIVVEKKRSINONNADIOYNS 119
QY	6 VDNFLLAANYDKRKLTHQGESIERFIKEGNQLPDERVIEERKRRLSTINTSDISTAT 65
Db	120 LASLTYPGALVKANSELVENOPDLVPYKRDSVTLSIDLPKVNHDEIIVQATKSIND 179
QY	66 NDSRLTPGALLVDETLENNPTLLADVRAPIYSIDLPGLASDSPLOYEDPSNSSVRG 125
Db	180 GVNTLVDMNNKRYSEHYENISAKIIDYDOEMAYSESOLAVKGAFAKNNSLANVFCAIS 239
QY	126 AVNDLLAKWHQDYGVN-NVPARROYEKTITHSMEOQLKAFGSPEFEKTSGLSIDERSVAH 184
Db	240 EGKVOEEVINRKQIIYYVNVNEPPSPSRFEKSVTKENLCALGYANAENPRAYSISVAAGR 299
QY	185 SGKOIQIVNKQIIYYTVSYDAVNNGPVDFODITVEDLKORGISABRPYITISSVAYGR 244
Db	300 DIFVLTSSHSSTRKAFAFDAGFKSGYKGDTLENIIONASFRAYVYGGSADAVEIID 359
QY	245 QVIKLKETTSSDEVEAAFEALINGVYAPATEMKQILDWTYKAVVLGGDPSSGARVYT 304
Db	360 GDLSKLRDIKOGANFNKRNKGPVIAATTNFLKMNQLAVVKNNSSEYIETTSKAYSQGIN 419
QY	305 GKVMVEDLLIOEGSRFTADHPGLPISTTSLRDYNVAFTQNSIDYEKTVATVIRNGDLL 364
Db	420 LDHGGAAYARNVNWDESYDANANEVEYKKSENDDCKLAHFHTSIYLPGAARINIH 479
QY	365 LDHGGAIVAAQIYTLTWNELSDHOEKVELTPRAMDRNQDILTAHFTTIDPLKGVNRILSVK 424
Db	480 AKECTGLAMEMWRTVVDNRNLPLVKNRNVCIWGTTLYPAYSDPYDN 525.
QY	425 IRECTGLAMEMWRTVEKTDLPVTRKRTISIKWGITLYPYVEDKVEN 470
RESULT	6
ENTRY	S22340 #type complete
TITLE	seeligeriolysin - Listeria seeligeri
ORGANISM	#formal name Listeria seeligeri
DATE	22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 21-Aug-1998
ACCESSIONS	S22340
REFERENCE	S22340

[illegible]

[illegible]

ORGANISM	#formal_name Bos primigenius taurus #common_name cattle
DATE	28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 29-Jan-1999
ACCESSIONS	S03974
REFERENCE	S03974
#authors	Powell, J.F.; Hsu, Y.P.; Weyler, W.; Chen, S.; Salech, J.;
#journal	Andrikopoulos, K.; Maliet, J.; Breakfield, X.O.
#title	Biochem. J. (1989) 259:407-413
#cross-references	The primary structure of bovine monoamine oxidase type A.
	Comparison with peptide sequences of bovine monoamine oxidase type B and other flavoenzymes.
#accession	MU02:89246344
#status	S03974
#molecule_type	not compared with conceptual translation
#residues	mRNA
#cross-references	1-527 #label POW
KEYWORDS	GB:X15609; NID:g5523; PID:g5524
	FAD; flavoprotein; mitochondrion; oxidoreductase
FEATURE	
15-43	
SUMMARY	
Query Match	3.8%; Score 124; DB 2; Length 527;
Best Local Similarity	20.0%; Pred. No. 1,95e-02;
Matches	30; Conservative 42; Mismatches 72; Indels 6; Gaps 5;
Db	1 MESLQKTSDAQMDVYVIGGIGISGLSAK-LIAEHVNVYLVLEARRVGGRTYVNEH 59
Oy	250 LETTSKSEVEPAEALIKGVKAPDTEMKOILDNTVTKAVILGGDSSGARVVTGKVDM 309
Db	60 V-DYVDVGAAYGVTQNRILRLSKOLELETYKVNVNERLHYVKGKTPPRGAFPPVWP 118
Oy	310 VEDLIQESRTADHPGLPISTYSF-LRNVVATFPQNSDIYETKTAARNQDLIDHS 368
Db	119 IAYL-DYNNLMRTM-DNMGKEIPADAPWE 145
Oy	369 GAYVAQYTYTWNELSYDHQKEVLETPKAMD 398
RESULT	13
ENTRY	T03415 #type complete
TITLE	S-layer protein - Campylobacter rectus
ORGANISM	#formal_name Campylobacter rectus
DATE	16-Mar-1999 #sequence_revision 16-Mar-1999 #text_change 16-Mar-1999
ACCESSIONS	T03415
REFERENCE	Z14942
#authors	Wang, B.; Kraig, E.; Kolodrubetz, D.
#journal	Infect. Immun. (1998) 66:1521-1526
#title	A new member of the S-layer protein family: characterization of the crs gene from Campylobacter rectus.
#accession	T03415
#status	preliminary
#residues	1-1361 #label MAN
#cross-references	EMBL:AF010143; NID:g2459960; PID:g2459961;
SUMMARY	#length 1361 #molecular_weight 144385 #checksum 6731
Query Match	3.6%; Score 118; DB 3; Length 1361;
Best Local Similarity	22.0%; Pred. No. 1.03e-01;
Matches	41; Conservative 49; Mismatches 86; Indels 10; Gaps 10;
Db	178 VDMHSNPHEPKAVELTTNDNATANFANPMK-HNPGGIDRIMTLQSSDKLTGDYSRHD 236
Oy	115 VEDPSNSSVRGAVNDLLAKWHODYGVNNVPARQYKEKITAHSME-OLKVKFGSDEKGT 173
Db	237 NTLVNEFGQANADGDPTRPTLTNTIONINIEYTGVTNLLD-RQSNDEKXINIRITK 295
Oy	174 NSLDIDNSVHSGE-KOI-QTVNEKQI-YTVSVDAVKKNQGDVFDQVTVEDLKQKGIS 230
Db	296 EAGNKFVNESIG-OKLVGMRLANAKKD-IDVKEEH-KKCVLSGFEDKRSVPLENFEAKS 352

QY 231 ERLPVY-ISSVAVGROYVLTETTSKSDVEAFALIKGVAVAQTEWKQILDVTEYKA 289
DB 353 LSTSP 358
QY 290 VILGGD 295

RESULT 14
ENTRY 560947 #type complete
TITLE hypothetical protein YOR220w - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES hypothetical protein O5050; hypothetical protein YOR50-10
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 05-Jun-1998
ACCESSIONS S60947; S67113; S71722
REFERENCE S60938
#authors Gallison, F.; Dujon, B.
#submission submitted to the EMBL Data Library, October 1995
#description Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of the yeast Saccharomyces cerevisiae.
#accession S60947
##molecule_type DNA
##residues 1-265 #label GAL
#cross-references EMBL:X92441; NID:g1050762; PID:g1050772
REFERENCE S67104
#authors Boyer, J.; Fairhead, C.; Gallion, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon, B.
#submission submitted to the Protein Sequence Database, July 1996
#accession S67113
##molecule_type DNA
##residues 1-265 #label BOY
#cross-references EMBL:Z75128; NID:g1420509; PID:e252085; PID:g1420510; MIRS:YOR220w
##experimental_source strain S288C
REFERENCE S71713
#authors Gallison, F.; Dujon, B.
#journal Yeast (1996) 12:877-885
#title Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of the yeast Saccharomyces cerevisiae.
#cross-references MUID:96437977
#accession S71722
##status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-265 #label GAM
#cross-references EMBL:X92441; NID:g1050762; PID:g1050772
#note the nucleotide sequence was submitted to the EMBL Data Library, October 1995

GENETICS
#map_position 15R
SUMMARY #length 265 #molecular_weight 29255 #checksum 6028
Query Match 3.5%; Score 117; DB 2; Length 265;
Best Local Similarity 27.6%; Pred. No. 1.35e-01;
Matches 29; Conservative 28; Mismatches 41; Indels 7; Gaps 7;

DB 85 PMKLFVETSLNNO-HPR-SRSTDVAISLONNALLDHRNRKLSINDPVGTVGDS 142
QY 12 AMN-YDKKKLLTHGSELENFIREGNOLPDE-FVYLE-RRKRSLSTNTSDISTATNDS 68
DB 143 SLNKGSSLSFDKSLSPMLKLSIDSPKPSYOEPLPKLSRSS 187
QY 69 RLY-PGALLVVDFTLLNNPTL-LAVDRAPMTYSIDLPGIASSDS 111

RESULT 15
ENTRY S58322 #type complete
TITLE nucleolar protein NOP5 - yeast (Saccharomyces cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 24-Sep-1998

ACCESSIONS S58322; S67216; S71990
REFERENCE S58318
#authors Pearson, B.M.; Hernando, Y.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.
#submission submitted to the EMBL Data Library, August 1995
#accession S58322
##molecule_type DNA
##residues 1-511 #label PEA
#cross-references EMBL:X90565; NID:g940836; PID:g940841
REFERENCE S67213
#authors Pearson, B.M.; Hernando, Y.; Kalogeropoulos, A.; Schweizer, M.
#submission submitted to the Protein Sequence Database, July 1996
#accession S67216
##molecule_type DNA
##residues 1-511 #label PEW
#cross-references EMBL:Z75217; NID:g1420680; PID:e252143; PID:g1420682; MIRS:YOR310C
##experimental_source strain S288C
REFERENCE S71986
#authors Pearson, B.M.; Hernando, Y.; Payne, J.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.
#journal Yeast (1996) 12:1021-1031
#title Sequencing of a 35.71 kb DNA segment on the right arm of chromosome XV reveals regions of similarity to chromosomes I and XIII.
#accession S71990
##status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-511 #label PER
#cross-references EMBL:X90565; NID:g940836; PID:g940841
#note the nucleotide sequence was submitted to the EMBL Data Library, August 1995

GENETICS
#gene SGD:NOPS
#cross-references SGD:S0005837; MIRS:YOR310C
FUNCTION #map_position 15R
#description involved in the synthesis of the 40S ribosomal subunit
KEYWORDS nucleus
SUMMARY #length 511 #molecular_weight 56956 #checksum 4899
Query Match 3.4%; Score 112; DB 2; Length 511;
Best Local Similarity 22.8%; Pred. No. 5.16e-01;
Matches 31; Conservative 39; Mismatches 58; Indels 8; Gaps 8;

DB 188 YGWH-FPEELAKIV-TDSVAVARIITLTMGRSKASPTDSEILPEIEER-VKTAAEVSMG 244
QY 242 YGRQVYLKLETTSKSDVEAFALIKGVK-APQTEWKQILDVTEYKAVILGGDPSSG 299
DB 245 TEITQTDLDNINALAEQIVFAAYREOLS-NYLSARMAKAIPNLITOLVGEIVGARLIA-H 302
QY 300 ARVYVGKVDVMDLEDLQEGSRFTADHGPLISTYTSFLDNVAVFQNSTDVEIKVAYR 359
DB 303 SGLSLSLAKSPASTIQ 318
QY 360 NGDLL-LDHSAGVVAQ 374

Search completed: Mon Aug 30 16:08:05 1999
Job time : 59 secs.

 WIRESETH
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 30 16:04:38 1999; Maspar time 14.67 Seconds

Tabular output not generated. 907.352 Million cell updates/sec

Title: >US-09-120-044-3

Description: (1-471) from US09120044.pep

Sequence: 1 MANKAVNDFILANMYDKKL.....TISWGTTLYPYQEVKEND 471

Scoring table: PAM 150
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot37
1:swissprot

Statistics: Mean 51.760; Variance 98.133; scale 0.527

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description	Pred. No.
1	3286	99.6	470	TACY_STRPN	0.00e+00
2	1618	49.0	500	TACY_CLOPE	0.00e+00
3	1558	47.2	529	TACY_LISMO	4.39e-303
4	1539	46.7	528	TACY_LISTIV	7.16e-299
5	1502	45.5	530	TACY_LISTE	1.13e-290
6	1492	45.2	501	TACY_PAEAL	1.85e-288
7	1437	43.6	574	TACY_STREQ	2.74e-276
8	1436	43.5	571	TACY_STRPY	4.56e-276
9	1430	43.0	574	TACY_STRCB	1.57e-272
10	1358	41.2	485	TACY_BACCE	7.75e-259
11	1344	3.8	527	AMINE OXIDASE [FLAVIN-NUCLEOLAR PROTEIN NOPS	1.09e-01
12	1112	3.4	511	NOPS_YEAST	1.97e-01
13	110	3.3	872	DP3A_MYCPN	1.97e-01
14	110	3.3	1228	SLAP_BACST	1.97e-01
15	105	3.2	146	CASK_UNGUN	6.42e-01
16	106	3.2	462	MPFB_YEAST	6.33e-01
17	102	3.1	418	Y285_HELPY	1.96e+00
18	101	3.1	427	HYPOHETICAL PROTEIN H	2.58e+00
19	103	3.1	512	ANX7_XENLA	1.48e+00
20	101	3.1	748	HRPH_PSESY	1.58e+00
21	102	3.1	756	V328_MYCGE	1.56e+00
22	101	3.1	1123	V120_HSV11	2.58e+00
23	98	3.0	210	SOMA_ONCKI	5.84e+00

24	99	3.0	235	1	REBP_CSV	4.46e+00
25	99	3.0	350	1	ARGI_SOYBN	4.46e+00
26	99	3.0	504	1	FLIC_SALMU	4.46e+00
27	98	3.0	513	1	YJIV_ECOLI	5.84e+00
28	100	3.0	523	1	NAM8_YEAST	3.40e+00
29	99	3.0	527	1	AOFA_HUMAN	4.46e+00
30	99	3.0	723	1	HS9B_MOUSE	4.46e+00
31	100	3.0	848	1	AMPN_LACTA	4.46e+00
32	99	3.0	993	1	VIA_CMT1	3.40e+00
33	98	3.0	1026	1	BGAL_STRTR	4.46e+00
34	95	2.9	177	1	GCSH_YEAST	5.84e+00
35	97	2.9	364	1	GCH2_PHOIE	7.63e+00
36	96	2.9	402	1	ARGJ_METJA	9.94e+00
37	97	2.9	419	1	Y223_MYCPN	9.94e+00
38	96	2.9	527	1	IF2G_YEAST	7.63e+00
39	96	2.9	662	1	LOXE_MOUSE	9.94e+00
40	97	2.9	722	1	YG42_YEAST	9.94e+00
41	96	2.9	767	1	HYPF_SYNY3	7.63e+00
42	97	2.9	1235	1	CYAD_TRYBB	9.94e+00
43	97	2.9	1390	1	RPOB_MYCGE	7.63e+00
44	97	2.9	1675	1	CLH_RAT	7.63e+00
45	96	2.9	3027	1	POLG_PYFV1	9.94e+00

ALIGNMENTS

RESULT 1
ID TACY_STRPN STANDARD: PRT: 470 AA.

AC P11990;
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE PNEUMOLYSIN (THIOL-ACTIVATED CYTOLYSIN).

GN PLY.
OS STREPTOCOCCUS PNEUMONIAE.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
CC STREPTOCOCCUS.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-NCC 7466 SEROTYPE II;
RX MEDLINE: 87193109.

RA WALKER J.A., ALLEN R.L., FALMAGNE P., JOHNSON M.K., BOUNOIS G.J.;
RT "molecular cloning, characterization, and complete nucleotide
sequence of the gene for pneumolysin, the sulfhydryl-activated toxin
of Streptococcus pneumoniae";

RL INFECT. IMMUN. 55:1184-1189(1987).

CC -1- FUNCTION: SULPHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.

CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL: X52474; G47404; -
CC EMBL: M17717; G153692; -
CC PIR: A28566; A28568.
CC PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
CC PFAM: PF01289; THIOL_cytolysin; 1.
CC HSSP: P19995; LPFO.
CC TOXIN; HEMOLYSIS; CYTOLYSIS.

FT INIT_MET 0
FT ACT_SITE 427
SQ SEQUENCE 470 AA; 52768 MW; D3F3A252 CRC32;

Query Match 99.6%; Score 3286; DB 1; Length 470;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;

Matches 469; Conservative 1; Mismatches 0; Indels 0; Gaps

Dd 1 ANKAVNDFLIANNMYCKKKLLTHQGSINIRIKRGNDLPDEFVIARRKRSLSTNTSDIS 60
|||
Oy 2 ANNAVNDFFLIANNMYCKKKLLTHQGSIEINRIKRGNDLPDEFVIARRKRSLSTNTSDIS 61
|||
Dd 61 VTATNDSRLYPGALLVPDETLLENPTLLADRAPMTYSIDLPGLASDSPFOVEDPNS`120
|||
Oy 62 VTAITNDSRLYPGALLVPDETLLENPTLLADRAPMTYSIDLPGLASDSPFOVEDPNS 121
|||
Dd 121 SVYGAVNDLAKHODYGVNNVPAARMQYEKITASHSEQLVKFGSDFEKTGNLIDFN 180
|||
Oy 122 SVYGAVNDLAKHODYGVNNVPAARMQEKITASHSEQLVKFGSDFEKTGNLIDFN 181
|||
Dd 181 SVSYSGEKOLIVNEKOIYYTVSVDAVKRNGVFDDTYVDELKORGISAERPLVYISSA 240
|||
Oy 182 SVSYSGEKOLIVNEKOIYYTVSVDAVKRNGVFDDTYVDELKORGISAERPLVYISSA 241
|||
Dd 241 YGRQVYLKETTSKSDEVEAEAFALIKGVAPOTEWKQILDNTEYKAVILGDSPSSGAR 300
|||
Oy 242 YGRQVYLKETTSKSDEVEAEAFALIKGVAPOTEWKQILDNTEYKAVILGDSPSSGAR 301
|||
Dd 301 VVGKVDWMDIELQESRRTAHPELPYSTTSLRNVAATONSTDYVERKYATARYNG 360
|||
Oy 302 VVGKVDWMDIELQESRRTAHPELPYSTTSLRNVAATONSTDYVERKYATARYNG 361
|||
Dd 361 DLLLDHSAGYVAQYYITWDELSDYDHQGEKVLTTPAMDRNQGLLAHTTISIPLKGNVRL 420
|||
Oy 362 DLLLDHSAGYVAQYYITWELSDYDHQGEKVLTTPAMDRNQGLLAHTTISIPLKGNVRL 421
|||
Dd 421 SVVIRECTGLAMEMARTYERKDLPVLYKKRTISWGTTILPOVEDKAYEND 470
|||
Oy 422 SVVIRECTGLAMEMARTYERKDLPVLYKKRTISWGTTILPOVEDKAYEND 471
|||

percent 2
ID TACY.CLOPE STANDARD: PRT: 500 AA.

AC p19995:
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DJ 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE PERFINCOLVISIN O PRECURSOR (THETA-TOXIN) (THIOU-ACTIVATED CYTOLYSIN).
GN PRO OR PEOR OR PECA.
OS CLOSTRIDIUM PERFINGENS.
OC PLASMID PT1B.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;
OC CLOSTRIDIUM.
CN [1]
RA SEQUENCE FROM N.A.
RX MEDLINE: 91099951.
RT SHIMIZU T., OKABE A., MINAMI J., HAYASHI H.;
RL "An upstream regulatory sequence stimulates expression of the
perfringolysin O gene of Clostridium perfringens.";
LN INFECT. IMMUN. 59:137-142(1991).
[2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-44.
RC STRAIN=ATCC 13124;
RE MEDLINE: 89932623.
RA TWETER R.K.;
RT "Nucleotide sequence of the gene for perfringolysin O (theta-toxin)
from Clostridium perfringens: significant homology with the genes for
streptolysin O and pneumolysin.";
RL INFECT. IMMUN. 56:3235-3240(1988).
[3]
RN SEQUENCE OF 29-45 AND 305-312.
RP STRAIN=P86K;
RC STRAIN=P86K;
RX MEDLINE: 87076517.
RA OHNO-IWASHITA Y., IWAMOTO M., MITSUI K., KAWASAKI H., ANDO S.;
RL "Cold-labile hemolysin produced by limited proteolysis of theta-toxin
from Clostridium perfringens";
RL BIOCHEMISTRY 25:6048-6053(1986).
[4]

```

RX STRAIN-NCIC 8237;
RY MEDLINE: 960123363
RA SHIMIZU T., KOHAYASHI T., BA-THEIN W., OHTANI K., HAYASHI H.;
RT "Sequence analysis of flanking regions of the pfoA gene of Clostridium perfringens: beta-galactosidase gene (pbq) is located in the 3'-flanking region.";
RL MICROBIOL. IMMUNOL. 39:677-686(1995).
RP [5]
RR IMPORTANCE OF THIOL-GROUP IN CHOLESTEROL BINDING.
RX MEDLINE: 88004463.
RA IMAMOTO M., OHNO-IWASITA Y., ANDO S.;
RT "Role of the essential thiol group in the thiol-activated cytolysin from Clostridium perfringens";
RN EUR. J. BIOCHEM. 167:425-430(1987).
RP [6]
RQ X-RAY CRYSTALLOGRAPHY (2.4 ANGSTOMS).
RX MEDLINE: 97113505.
RA FEIL S.C., ROSSJOHN J., RODE K., TWETEN R.K., PARKER M.W.;
RT "Crystallization and preliminary x-ray analysis of a thiol-activated cytolysin.";
RN FEBS LETT. 397:290-292(1996).
RP [7]
RQ X-RAY CRYSTALLOGRAPHY (2.2 ANGSTOMS).
RX MEDLINE: 97325744.
RA ROSSJOHN J., FEIL S.C., MCKINSTREY W.J., TWETEN R.K., PARKER M.W.;
RT "Structure of a cholesterol-binding, thiol-activated cytotoxin and a model of its membrane form.";
RN CELL 89:685-692(1997).
CC -! FUNCTION: SUPERNATANT-ACTIVATED TOXIN IS ABLE TO LYSO CHOLESTEROL CONTAINING MEMBRANES CAN BE REVERSIBLY INACTIVATED BY OXYDATION. CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO EUKARYOTIC CELL MEMBRANES.
CC -! SUBUNIT FORMS OLIGOMERS IN THE HOST MEMBRANE.
CC -! SIMILARITY BELONGS TO THE THIOLO-ACTIVATED CYTOLYSIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/cc or send an email to license@isb-sib.ch).
-----
DR EMBL: M81080; GI44886; -.
DR EMBL: M65704; GI44884; -.
DR EMBL: D49537; G1502275; -.
DR PIR: B43577; B43577.
DR PDN: LPFO: 05-AUG-98.
DR PROSITE: PS00481; THIOLO-CYTOLYSINS; 1.
DR PFAM: PF01289; ThioLO CytoLySiNs; 1.
FM TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMID; 3D-STRUCTURE.
FT SIGNAL
   1      28
FT CHAIN
   29     500
FT ACT_SITE
   459    459          PERFRINGOLYSIN O.
CONFLICT       30        30         D -> K (IN REF. 3 ).
CONFLICT       34        34         K -> I (IN REF. 3 ).
CONFLICT       126       128         RKP -> EA (IN REF. 2 ).
FT SEQUENCE
   500 AA; 55799 MW; 17FC32BC CRC32;.
SQ
Query Match           49.0%; Score 1618; DB 1; Length 500;
Best Local Similarity 46.9%; Pred. No. 0.0e+00;
Matches 216; Conservative 110; Mismatches 132; Indels 3; Gaps 3;
```

QY	122	SVRCGAVNLLAKKHODDYQGVANNVPARMQYKIKTAHSMEQLKVKRGDSPEKTKNSLDIDEN	161
Db	213	AVANNEKVMILAAKQIPIYTSADLPKRPSPDLFPDSDVTFNDLKQKGYNSNEAPPLMVSNA	272
QY	182	SVHSEKQIQIYNEKQIITYIVSVPAAKMGVFPDDTYVEDLKRGSIABRPPLYISSVA	244
Db	273	YGRPIYKALETTSSSKIDVOAFAKLIKNTDLKNSOOQYKDIYENSFPYAVVLGDAOEHNK	332
QY	242	YGRQYVLETTETSSDEVEAEFAELIKGVKVAPOTEWKQILDMTEYKAVILIGOPSSGAR	300
Db	333	VYTKDFEELIRKIVKDNAPFSTKKNAPRPISTYSVFLKONSVAAYHANKDIYIETTISTEYSG	392
QY	302	VYTGKGVNVEDLIQEGSFYIDHFGPLPSTYTSFLRDNVAVATFQNSUDYVETKRYATARG	361
Db	393	KINLDHSGAYVAQEVAMDEVSYDKENENYLTHTKTWGDQNDKTAHSTYIPELANARI	452
QY	362	DLLLDHSGAYVAQYIITWNELSYDHQGEVLTTPKAMPBNGODTLTAHFTTSIPLKGNVNL	421
Db	453	RIRARECTGLAMEMMRDVIISDYDPLNNINVSITWGTTIYR	493
QY	422	SVKTRRECTGLAMEMMRVITYEKTDLPVPRKRTISITWGTTIYR	462

```

RESULT      3
ID          3          STANDARD:          PRT:          529 AA.
AC          P13128: 048747.; 057096; 057206;
DT          01-JAN-1990 (REL. 13, CREATED)
DT          01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT          15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE          LISTERIOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
HAY OR LISA OR HLXA.
OS          LISTERIA MONOCYTOGENES.
OC          BACTERIA, FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
RN          LISTERIA.
RP          [1]
RC          SEQUENCE FROM N.A.
RX          STRAIN-EGD / SEROTYPE 1/2A;
RX          MEDLINE: 89366684.
RA          DOMANN E., CHAKRABORTY T.;
RT          "Nucleotide sequence of the listeriolysin gene from a Listeria
RN          monocytogenic serotype 1/2a strain."
RL          NUCLEIC ACIDS RES. 17:6406-6406(1989).
RX          [2]
RC          SEQUENCE FROM N.A.
RX          MEDLINE: 88153053.
RA          MENGAUD J., VICENTE M.-F., CHENEVERT J., PEREIRA J.M., GEOFFROY C.,
RA          GICQUEL-SANZET B., BAQUERO F., PEREZ-DIAZ J.-C., COSSART P.;
RN          "Expression in Escherichia coli and sequence analysis of the
RL          listeriolysin O determinant of Listeria monocytogenes."
RX          INFECT. IMMUN. 56:766-772(1988).
RN          [3]
RC          SEQUENCE FROM N.A.
RX          STRAIN-12067;
RX          MEDLINE: 92040062.
RA          RASMUSSEN O.F., BECK T., OLSEN J.E., DONS L., ROSSEN L.;
RN          "Listeria monocytogenes isolates can be classified into two major
RT          types according to the sequence of the listeriolysin gene."
PL          INFECT. IMMUN. 59:3945-3951(1991).
RN          [4]
RC          SEQUENCE FROM N.A.
RX          STRAIN-5782 (SEROTYPE 4B), AND F4233, F6789 (SEROTYPE 1/2B);
RN          VINES A., SWAMINATHAN B.;
RN          SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC          -1- FUNCTION: SUDHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSO CHOLESTEROL
CC          CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC          CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC          EUKARYOTIC CELL MEMBRANES.
CC          -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC          THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC          the European Bioinformatics Institute. There are no restrictions on lists
CC          use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR	EMBL	A13127	G44107	-
DR	EMBL	M24199	G149653	-
DR	EMBL	X60035	G44112	-
DR	EMBL	U25452	G887870	-
DR	EMBL	U25443	G887016	-
DR	EMBL	U25446	G887028	ALT_INIT.
DR	EMBL	U25449	G887864	ALT_INIT.
DR	PIR	S05306	S05306	
DR	PIR	A43505	A43505	
DR	PROSITE	PS00481	THIOL_CYTOLYSIS	1.
DR	PFAM	PF01289	Thiol_cytolysis	1.
DR	HSSP	P19985	1PRO.	
KM	TOXIN	HEMOLYSIS	CYTOLYSIS	SIGNAL.
FT	SIGNAL	1	25	
FT	CHAIN	26	529	
FT	ACT_SITE	444	484	
FT	VARIANT	35	35	
FT	VARIANT	39	39	
FT	VARIANT	438	438	
FT	VARIANT	523	523	
FT	SEQUENCE	529 AA	58688 MW	269EA737 CRC32

Query Match	47.28;	Score 1558;	DB 1;	Length 529;
Best Local Similarity	43.38;	Pred. No. 4.39e-303;		
Matches	202;	Conservative 120;	Mismatches 143;	Indels 1;
			Gaps	1

ID	TACT	LIST	STANDARD	PRT	528 AA
DB	61	IKDYIOGDIDNNNNLYHGCAVAVNNP	RRKGYDGEYIVVKKKKKSI	NNONNADIQVNA	120
QY	6	VNDFILAMNYDKKLLIHQGESIENR	IKKGNOLPDEFVIERKKRS	LSTVTSIDISVAT	65
DB	121	ISLSAYPGALYKANSELVEONPDVL	PAVRDLSLSDIDEGMTNOON	LYVKKATKSVNN	180
QY	66	NNSRLYPGALLVDETLENNP	PLTIANVDRAPMTYSIDIE	PGLASSSFLQVDEDPNSVYG	125
DB	181	AVNTLVKMERNEKYAQAVPNVSAK	IDYDEMAVSESOLIAKEGTA	KAVANNSLANFGAIS	240
QY	126	AVNDLLAKWHDYQGVN	VNPARQWYEKTAHSMQEL	KYKFSQDEKGNLDDIDFNSVH	184
DB	241	EKKMEDEYISKQIYYVANNNEP	RRSFRFGKAVYKEQLQAL	GVNAENPRAYISSVAYGR	300
QY	185	SEKQIOIYNNKQIYYVSDAVK	NGDVFODIYVEDLKRGISA	ERPLVYISSVAYGR	244
DB	301	OYVLKLTNSHSTKYAKAPAAVS	GKSVSDVDLNIINKNSF	KAVIYGSAAKEVOIID	360
QY	245	OYVLKLETTSSKDEEAAFE	LKGVKVAQOETWKOI	LDNTEVKKVILIGDPSGAAVVT	304
DB	361	GNLGDLRDILKKGATFNERT	PGVPIATNPLKDNELAV	KNNSEYETTSKATDOKIN	420
QY	305	GKVDVADDELIOGSRFPA	DRHGLPISTTISFLRD	VNVAFPQSDTYETKTAARNDL	364
DB	421	IDHSGGYAQENISMDENVAND	PEGNEIYVHKWMSSENN	SKLAHFSTSYIEPGANRINIV	480
QY	365	LDHSGYAAQZYIIMNELSL	YDQKQEVLPKAMDNR	GODLANHFTTSIPLKGNVRLSVK	424
DB	481	AACETGLGLEMWRKVIDDR	RLPLVKNRNTSINGTLL	YPRYSKNYDN	526
QY	425	IRECTGLAMEWNRIVYEK	TDPLVRKRTISINGTLL	YQVEDKYEN	470
RESULT	4				
AC	P31831				
DT	01-JUL-1993	(REL. 26, CREATED)			
DT	01-JUL-1993	(REL. 26, LAST SEQUENCE UPDATE)			
DT	15-JUL-1998	(REL. 36, LAST ANNOTATION UPDATE)			
DE	IVANOLYSIN	PREDICTOR (THIO-L-ACTIVATED CYTOLYSIN).			

[illegible]

ID	RESULT	5	STANDARD;	PRT;	530 AA.
AC	TACY_LISSE				
AC	P1830;				
DT	01-JUL-1993 (REL. 26, CREATED)				
DT	01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)				
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)				
DE	SEELIGERIOLOYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).				
GN	L50.				
OS	LISTERIA SEELIGERI.				
OC	BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;				
OC	LISTERIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SLCC;				
RX	MEDLINE; 92182018.				
RA	HAAS A., DOMBSKY M., KREFT J.:				
RT	"listeriolysin genes: complete sequence of llo from <i>Listeria ivanovii</i>				
RL	and of lso from <i>Listeria seeligeri</i> ."				
RL	BIOCHIM. BIOPHYS. ACTA 1130:81-84(1992).				
CC	-1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSO CHOLESTEROL.				
CC	CONTAINING MEMBRANES. CAN BE REVERSELY INACTIVATED BY OXIDATION.				
CC	CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO				
CC	EUKARYOTIC CELL MEMBRANES.				
CC	-1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaborat				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commerci				
CC	entities requires a license agreement (See http://www.lsb-sib.ch/announce				
CC	or send an email to license@lsb-sib.ch).				
CC	-----				
DR	EMBL; G60462; G44145; -				
DR	PIR; S22340; S22340.				
DR	PROSITE; PS00481; THIOL_CYTOLYSIN; 1.				
DR	PFAM; PF01289; Thiol_cytolysin; 1.				
DR	HSSP; P19995; 1PFO.				
KW	TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.				
FT	SIGNAL	1	25		
FT	CHAIN	26	530	SEELIGERIOLOYSIN.	
FT	ACR SITE	485	485	BINDING TO CHOLESTEROL (BY SIMILARITY).	
SEQ	SEQUENCE	530 AA;	59181 MW;	68A497A5 CRC32;	
Query Match		45.5%;	Score 1502;	DB 1;	Length 530;
Best Local Similarity		43.1%;	Pred. No.1,13e-290;		
Matches	201; Conservative	115;	Mismatches 149;	Indels	1; Gaps 1.
Db	62 INKTYMGNTYGNKNSLTVYOGCAVYNNVPKKYKGGSEIYVKKKKGINONNADISYINA	121			
Qy	6 VNDFLANVYKAKKLLTHQGESINRFKIBENQJLPDEFVYIERKKRSLSITSDISTAT	65			
Db	122 ISSLTYPGALYKANRELVENOPNVLPAVKRDSLISVDLPGMTKDKNKIFVKNPFXSVNN	181			
Qy	66 NDSRLYPGALLVVDLELLNPNLLADVADRAPIYSIDLPGIASSDSFLQVEDDPNSSEVRG	125			
Db	182 AVNTLVEWMNKYSKAYENINAKIDISDEMAKYSQSOLAKGTAFKAVNNSLANFEAS	241			
Qy	126 AVNDLLAKMHODYGVN-NVPARMQYEKTIHSHMQLKVKYKSGPFEKTKGNSLIDFN	184			
Db	242 DKGVDVEVIFSRQIYNNINNEPTSPKFFGGSYTKEQDLALGVNAENPPAYISSVAYGR	301			
Qy	185 SGEKQIIVNFKQIYTVYVAVKMPGVDVQDTVEDLKQKQISAEERPLVYISSVAYGR	244			
Db	302 QYVYVALLSSSHSNKKVTKFEAAAGSKSVKGVDELNIIKKNSFFRAVLYIGSGAKREVEI	361			
Qy	245 QYVLLLETTSSDEVEAFELALIGVKAAPDETKMQLIDNTEVAAVILGDPSSGARVYT	304			
Db	362 GNIGELRLIKKSGSYDRENGVPISTYTNLKLNDLAVVNNSESEIETTSKSYTDKIN	421			
Qy	305 GKVDVVEDLLIDGSGSFYADHPGLPISTYTSFLRDNVAVATPQNSIDYETKVTAYRNDGL	364			
Db	422 IDHSGYVAQENISMDVEYSIDENGKELVKKHKGWENIKSLAHTTSSTLYLPGNARNIN	481			


```

QY 365 LDHSGAYVAGYITWNLSDYHOGKEVLTLPKAWDRNGODLTAHFTTSIDPLKGVNRLSVK 424
Db 482 ARCTGJFEMWRTYIDRLPLVKNRNVISIMCTTIXPHSNVDN 527
QY 425 IRECTGLAMEWRTYVETDPLVRRRTISIMGTLTYPOVEDKVEN 470

RESULT 6
AC TACY PAEAL STANDARD: PRT: 501 AA.
AC P23564:
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ALVEOLYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN ALV.
OS PAENIBACILLUS ALVEI (BACILLUS ALVEI).
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC PAENIBACILLUS.
OC [1]
OC SEQUENCE FROM N.A.
OC MEDLINE: 91072294.
OC RA GEOFFROY C., MENGAUD J., ALOUF J.E., COSSART P.;
OC "Alveolysin, the thiol-activated toxin of Bacillus alvei, is
OC homologous to listeriolysin O, perfringolysin O, pneumolysin, and
OC streptolysin O and contains a single cysteine."
OC RT J. BACTERIOL. 172:7301-7305(1990).
OC [2]
OC SEQUENCE OF 33-60.
OC RA ALOUF J.E., GEOFFROY C., GILLES A.M., PALMAGNE P.;
OC (IN RAPPOUILLI R., ALOUF J.E., PALMAGNE P. (EDS.);
OC BACTERIAL PROTEIN TOXINS, PP.49-50, GUSTAV FISHER VERLAG,
OC STUTTGART (1990).
OC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
OC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
OC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
OC EUKARYOTIC CELL MEMBRANES.
OC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
OC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
OC between the Swiss Institute of Bioinformatics and the EMBL outstation -
OC the European Bioinformatics Institute. There are no restrictions on its
OC use by non-profit institutions as long as its content is in no way
OC modified and this statement is not removed. Usage by and for commercial
OC entities requires a license agreement (See http://www.isb-sib.ch/announce/
OC or send an email to license@isb-sib.ch).
OC CC
OC CC EMBL: M62709; GI42473; -.
OC DR PIR: A37858; A37858.
OC DR PROSITE: PS00481; THIOL-CYTOLYSINS; 1.
OC DR PFAM: PF01289; Thiol_cytolysin; 1.
OC DR HSSP: P19995; lPFO.
OC KM TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
OC FT SIGNAL 32
OC FT CHAIN 33
OC FT ACT_SITE 461 501 ALVEOLYSIN.
OC FT ACT_SITE 461 461 BINDING TO CHOLESTEROL (BY SIMILARITY).
OC FT CONFLICT 56 56 A -> L (IN REF. 2).
OC FT SEQUENCE 501 AA; 55268 MW; 80P2C94A CRC32;

Query Match 45.2%; Score 1492; DB 1; Length 501;
Best Local Similarity 42.3%; Pred. No. 1,85e-288;
Matches 192; Conservative 127; Mismatches 132; Indels 3; Gaps 3;

```

```

QY 129 DLAKWHDYGVQVNNVARRQYEKITAHSMEQLKVFSGSPFETGNSLIDFNSVSGEK 188
Db 222 KVMVAATKQIFVTYASGLPNNPSOLFDDSVTFEALARKGVSNAPPLMNSVAAGRTIYV 281
QY 189 QIOIVNFKQIYVTVDAVKNPGDVFQDTVYEDLKRGISABRPLVYISSVAGQVYL 248
Db 282 KLETTSKNDYQVTFKLLNPPSTQASGQYKDIYENSSFAVYVGGDAQTHNOVYTKDEN 341
QY 249 KLETTSSDVEAEFAELLIGKVAAPQTEWKQILDNTYKAVILGGDPSGARGVYTKVD 308
Db 342 VIOSVIRKDNAQFSSKNPAPYPISTYSVFLKDNSTAAVHNNTYEIETETTESKGIKLDH 401
QY 309 MVEDLIOGSRFTDHDGLPISTTSLRNVAATONSIDYETKVTAYRNGDLDLDS 368
Db 402 GAYVAQFEVYWDFFSDADQGEIYTRKSDWGNRDSARHFEPIPLPPAKNIRIFAREC 461
QY 369 GAYVAQYIYTWNLSDYHOGKEVLTLPKAWDRNGODLTAHFTTSIDPLKGVNRLSVIRGC 428
Db 462 TGLAMEWRTYVDEYNVPLASDINVSIMGTLTYP 495
QY 429 TGLAMEWRTYVETDPLVRRRTISIMGTLTYP 462

RESULT 7
AC TACY STREQ STANDARD: PRT: 574 AA.
AC O54114:
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS EQUISIMILIS.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
OC [1]
OC SEQUENCE FROM N.A.
OC RA STRAIN-SIMD-1;
OC RX MEDLINE: 95102113.
OC RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
OC IKARASHI H., YOTSUDO T.;
OC "Cloning and sequencing the streptolysin O genes of group C and group
OC RT G streptococci."
OC RL DNA SEQ. 4:325-328(1994).
OC CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
OC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
OC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
OC EUKARYOTIC CELL MEMBRANES.
OC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
OC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
OC between the Swiss Institute of Bioinformatics and the EMBL outstation -
OC the European Bioinformatics Institute. There are no restrictions on its
OC use by non-profit institutions as long as its content is in no way
OC modified and this statement is not removed. Usage by and for commercial
OC entities requires a license agreement (See http://www.isb-sib.ch/announce/
OC or send an email to license@isb-sib.ch).
OC CC
OC CC EMBL: D16824; G498301; -.
OC DR PROSITE: PS00481; THIOL-CYTOLYSINS; 1.
OC DR PFAM: PF01289; Thiol_cytolysin; 1.
OC DR HSSP: P19995; lPFO.
OC KM TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
OC FT SIGNAL 36
OC FT CHAIN 37
OC FT ACT_SITE 533 574 STREPTOLYSIN O.
OC FT ACT_SITE 533 533 BINDING TO CHOLESTEROL (BY SIMILARITY).
OC FT SEQUENCE 574 AA; 63991 MW; 9BBF31B4 CRC32;

Query Match 43.6%; Score 1437; DB 1; Length 574;
Best Local Similarity 42.4%; Pred. No. 2.74e-276;
Matches 194; Conservative 110; Mismatches 151; Indels 3; Gaps 3;

```

```

OY 6 VNFILAMNYDKKLLTHOGESIENRFI-KEGNOLPDEFVIERKKRSISTNTSDISVTA 64
DB 171 SVTDRTYPALOLANKGFTEKNPDVAVTNRKPNOKIHIDLPKMGDKAT-VEVNDPTRYANVS 229
OY 65 TNDRIYPPGALLVDETLLENPTLLAVDAPRTYSIDLPGLAASSDSFLQVDPSSSVR 124
DB 230 TAIIDLNVOMHNDYSGNTLPARTOYTESMYSKSOIEALNVNSKILDTGLIDFKSIS 289
OY 125 GAVNDLLAKKHODYGOVNNVPARMQYEKITAHSMEQLKXFGSDPEKIGNSLIDIDFNSVH 184
DB 290 KGEKVMIAAYKOIFTVSANLPNNADYFDSKVTFLKQKRGVSNAPPLFVSNVAGR 349
OY 185 SGEKOIIVNFKOIYTVSVDAVKNPGDVFDOTVVEDLKQGISAEPLVYISSVAYGR 244
DB 350 TVFKLETSSKSDVEAFSAALKGTDVKTNGKYSIDLENSFTAVVLGGDAEHNKVT 409
OY 245 QVYLKLETSSKSDVEAFSAALKGTDVKTNGKYSIDLENSFTAVVLGGDAEHNKVT 304
DB 410 KDFVIRNVIKDNATSRKNPAPISYTSVFLKNNKIAGVNNREYVETTSYTSKIN 469
OY 305 GKVDMEVDDLQESRFTADHPGLPISTYTSFLRDNVVATFONSTDVETKVTAYRNGDLL 364
DB 470 LSHOGAVVAQYEIMDEINVDKGEVITKRRDNNWYSKTSFSTVPLGANSRIRIM 529
OY 365 LDHSGAVVAQYITWMLSDYHOGKEVLPKANDRNGODLTAFTTSLDKGNVRLSVK 424
DB 530 ARECTGLAMEMWRIYKTDLPVLRKRTISIMGTILXP 567
OY 425 IRECTGLAMEMWRIYKTDLPVLRKRTISIMGTILXP 462

```

RESULT 8

ID TACY_STRPY STANDARD: PRT: 571 AA.

AC P21131;
01-FEB-1991 (REL. 17, CREATED)
01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
SLO.
STREPTOCOCCUS PYOGENES.
OG PLASMID PM157.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN SEQUENCE FROM N.A.
RX MEDLINE; 88057628.
RA KEHOE M.A., MILLER L., WALKER J.A., BOULNOIS G.J.;
RT "Nucleotide sequence of the streptolysin O (SLO) gene: structural
homologies between SLO and other membrane-damaging, thiol-activated
toxins.";
RT INFECT. IMMUN. 55:3228-3232(1987).
RL -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M18638; G153811; -
CC PIR; A43507; A43507.
CC PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
CC FRAM; PF01289; ThiOL_cytolysin; 1.
CC HSSP; P19995; 1PFO.
CC TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMID.
KM SIGNAL
FT CHAIN 1 33
34 571 STREPTOLYSIN O.

```

FT ACT SITE 530 530 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 571 AA; 63638 MW; 33124B54 CRC32;
OY 6 VNFILAMNYDKKLLTHOGESIENRFI-KEGNOLPDEFVIERKKRSISTNTSDISVTA 64
DB 109 INDIKISLVNNELEVLAKNGETIEN-FVBEKGVKADKFEIVIERKKNNINTPPVDISITD 167
OY 6 VNFILAMNYDKKLLTHOGESIENRFI-KEGNOLPDEFVIERKKRSISTNTSDISVTA 64
DB 168 SVTDRTYPALOLANKGFTEKNPDVAVTNRKPNOKIHIDLPKMGDKAT-VEVNDPTRYANVS 226
OY 65 TNDRIYPPGALLVDETLLENPTLLAVDAPRTYSIDLPGLAASSDSFLQVDPSSSVR 124
DB 227 TAIIDLNVOMHNDYSGNTLPARTOYTESMYSKSOIEALNVNSKILDTGLIDFKSIS 286
OY 125 GAVNDLLAKKHODYGOVNNVPARMQYEKITAHSMEQLKXFGSDPEKIGNSLIDIDFNSVH 184
DB 287 KGEKVMIAAYKOIFTVSANLPNNADYFDSKVTFLKQKRGVSNAPPLFVSNVAGR 346
OY 185 SGEKOIIVNFKOIYTVSVDAVKNPGDVFDOTVVEDLKQGISAEPLVYISSVAYGR 244
DB 347 TVFKLETSSKSDVEAFSAALKGTDVKTNGKYSIDLENSFTAVVLGGDAEHNKVT 406
OY 245 QVYLKLETSSKSDVEAFSAALKGTDVKTNGKYSIDLENSFTAVVLGGDAEHNKVT 304
DB 407 KDFVIRNVIKDNATSRKNPAPISYTSVFLKNNKIAGVNNREYVETTSYTSKIN 466
OY 305 GKVDMEVDDLQESRFTADHPGLPISTYTSFLRDNVVATFONSTDVETKVTAYRNGDLL 364
DB 467 LSHOGAVVAQYEIMDEINVDKGEVITKRRDNNWYSKTSFSTVPLGANSRIRIM 526
OY 365 LDHSGAVVAQYITWMLSDYHOGKEVLPKANDRNGODLTAFTTSLDKGNVRLSVK 424
DB 527 ARECTGLAMEMWRIYKTDLPVLRKRTISIMGTILXP 564
OY 425 IRECTGLAMEMWRIYKTDLPVLRKRTISIMGTILXP 462

```

RESULT 9

ID TACY_STRCB STANDARD: PRT: 574 AA.

AC 053957;
15-JUL-1998 (REL. 36, CREATED)
15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
SLO.
STREPTOCOCCUS CANIS.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN SEQUENCE FROM N.A.
RX MEDLINE; 95102113.
RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
RT "Cloning and sequencing the streptolysin O genes of group C and group
RT G streptococci.";
RL DNA SDO. 4:325-328(1994).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL: D16825; G498378; -
 DR PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
 DR PFAM: PF01289; Thiol_cytolysin; 1.
 DR HSSP: P19995; 1PFO.
 KM TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
 FT SIGNAL 1 36 BY SIMILARITY.
 FT CHAIN 37 574 STREPTOLYSIN O.
 FT ACT_SITE 533 533 BINDING TO CHOLESTEROL.
 SO SEQUENCE 574 AA; 63910 MW; 98BD1743 CRC32;
 Query Match 43.08; Score 1420; DB 1; Length 574;
 Best Local Similarity 42.1%; Pred. No. 1,57e-272;
 Matches 193; Conservative 108; Mismatches 154; Indels 3; Gaps 3;
 Db 112 INKITSLVNLEVLAKNETTEN-FVPEKGYKAKDKFVIERKKNTPTVDISTID 170
 QY 6 VNDPILAMNYDKKKLLTHOGESIENRFI-KEGNQLPDEFVIERKKRSLSTSDISVTA 64
 Db 171 SVTDRTYPAALQANKGFTENKPDVAVTRKPKIHIDPGMGDKAT-VEVNDPTAYNS 229
 QY 65 TNSRLYPGALLVYDETLLENPTLLAVDRAPMTYSIDLPGLASSDSFLQVEDPSNSVTR 124
 Db 230 TADNLYVNMHNDYSGCNTLPARTQYTESMYSKSQIEALNVNSKILDTGLGIDFKSIS 289
 QY 125 GAVNDLLAKWHQDYGVNNVPAMQYKELTAHSMQKVKFGSDEFKTSGLDIDENSVAH 184
 Db 290 KGEKRWIAAYKQIETVSANLPPNPADYEDSKVTFKELQAGVSNAPPLVYSNVAIGR 349
 QY 185 SGKKQIOLVNFQIYTVSDAVKNGDVFODTVYEDLKORGISERPLVYISSVAYGR 244
 Db 350 TVFKLETSSKNDVFAASALKGTVTKNGKYSIDLENSFTAVLADAAEHKKVYT 409
 QY 245 QYALKETTSKSDVEFAAEALIKGVKAPQTEMKOILDTEVKAVILGDSGSGARVYT 304
 Db 410 KQPDVIRNVIKANATSKRNPAVPISTYVFLKNNKIAGVNNRSEVETSTETYSKGIN 469
 QY 305 GKVDWEDLIDEGSRFTADHPGPISTYTSFLDNVAFQNSTDYVETKTVARNGDGL 364
 Db 470 LSHQGVVAYOYELTMDLNEVVDKGEVITKRRDNWYKTSFESVITPLGANSRIRIM 529
 QY 365 LHDHSGAYVAYOYITMELSYDHQGEVILTPKAMDNGODLTAHFTTSPILKGNVRLSVK 424
 Db 530 ARECTGLAEMWRTIVDERDVLSKEININISGTLSP 567
 QY 425 IRECTGLAEMWRTIVDERDVLSKEININISGTLSP 462
 RESULT 10
 ID TACY_BACCE STANDARD; PRT; 485 AA.
 AC 045105;
 DT 15-JUL-1998 (REL. 36, CREATED)
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 CC BACILLUS CEREUS.
 CC BACILLUS CEREUS.
 CC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 CC BACILLUS.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN-RIMD 206001;
 CC YUTSUDO T.;
 CC SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: SUFFHYDRI-ACTIVATED TOXIN. IS ABLE TO LYSO CHOLESTEROL
 CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
 CC EUKARYOTIC CELL MEMBRANES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D21270; G600252; -
 DR PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
 DR PFAM: PF01289; Thiol_cytolysin; 1.
 DR HSSP: P19995; 1PFO.
 KM TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 >485 HEMOLYSIN.
 FT ACT_SITE 466 466 BINDING TO CHOLESTEROL (BY SIMILARITY).
 FT NON_TER 485 485
 SO SEQUENCE 485 AA; 53862 MW; 45208ABA CRC32;
 Query Match 41.2%; Score 1358; DB 1; Length 485;
 Best Local Similarity 40.1%; Pred. No. 7,75e-259;
 Matches 175; Conservative 121; Mismatches 137; Indels 3; Gaps 3;
 Db 44 SIDGIGMLTYNNOEVLAVNGDKVYES-FVPEKESINSNGKRVVVDKRNHLQRRQSIFFRL 102
 QY 5 AVNDFILAMNYDKKKLLTHOGESIENRFI-KEGNQLPDEFVIERKKRSLSTSDISVTA 63
 Db 103 DSVANRTYPPGAVOLANKAPADNOPSLVAKRKPLNTSIDPGMKRNT-ITVQNPYGVN 161
 QY 64 ATNDSRLYPGALLVYDETLLENPTLLAVDRAPMTYSIDLPGLASSDSFLQVEDPSNSV 123
 Db 162 AGAVDDLVSTWNEKYSATHTLPARMQYTESMYSKAQIASALNVNKKYLDNSINIPFNAY 221
 QY 124 RGAVNDLLAKWHQDYGVNNVPAMQYKELTAHSMQKVKFGSDEFKTSGLDIDENSVAH 183
 QY 222 ANGKRWIAAYKQIETVSANLPPNPADYEDSKVTFKELQAGVSNAPPLVYSNVAIGR 281
 QY 184 HSGEKQIOLVNFQIYTVSDAVKNGDVFODTVYEDLKORGISERPLVYISSVAYGR 243
 Db 282 RTVYVKLETSSKNDVFAASALKGTVTKNGKYSIDLENSFTAVLADAAEHKKVYT 341
 QY 244 ROYALKETTSKSDVEFAAEALIKGVKAPQTEMKOILDTEVKAVILGDSGSGARVYT 303
 Db 342 TKDNEIRNIITKDAELSFKNPAVPISTYVFLKNNKIAGVNNRSEVETSTETYSKGIN 401
 QY 304 TGKVDWEDLIDEGSRFTADHPGPISTYTSFLDNVAFQNSTDYVETKTVARNGDGL 363
 Db 402 TLHDHGAIVAYOYDVSMDGFTFDNGKEILTKTWESGDKDTAHSYTVLPPLPNSKNIKI 461
 QY 364 LLDHSGAYVAYOYITMELSYDHQGEVILTPKAMDNGODLTAHFTTSPILKGNVRLSVK 423
 Db 462 VARECTGLAEMWRTIV 477
 QY 424 KIRECTGLAEMWRTIV 439
 RESULT 11
 ID AOPA_BOVIN STANDARD; PRT; 527 AA.
 AC P21398;
 DT 01-MAY-1991 (REL. 18, CREATED)
 DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE AMINE OXIDASE [FLAVIN-CONTAINING] A (EC 1.4.3.4) (MONOAMINE OXIDASE)
 GN MAOA.
 CC BOS TAURUS (BOVINE).
 CC EDAROTIA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC ARIODACTYLIA; RUMINANTIA; PECORA; BOVOIDEA; BOVIDAE; BOVINAE; BOS.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE: 89246344.
 CC POWELL J.F., HSU Y.P.P., WEYLER W., CHEN S., SALACH J.,
 CC ANDRIKOPOULOS K., MALLETT J., BREAKFIELD X.O.;
 CC "The primary structure of bovine monamine oxidase type A. Comparison
 CC with peptide sequences of bovine monamine oxidase type B and other
 CC flavoenzymes.";
 CC RT BIOCHEM. J. 259:407-413(1989).

CC EMBL: AE000045: G1674148: -
 DR TRANSFRASE: DNA-DIRECTED DNA POLYMERASE: DNA REPLICATION.
 KW SEQUENCE 872 AA: 99257 MW: C6C0C49B CRC32;
 SQ

Query Match 3.3%; Score 110; DB 1; Length 872;
 Best Local Similarity 25.4%; Pred. No. 1.97e-01;
 Matches 32; Conservative 29; Mismatches 57; Indels 8; Gaps 8;

DB 302 FAHNNDLILGPGSGAVSLA-FLIKITQIDPVANNLFEFFSRHQGLP-DIDIDIM 359
 QY 112 FLQVEDPSSSVRG-AYNDLAKMHQDYGQVNNVPARQYKETAHSQKFKGSDFE 170
 DB 360 ETKRDLVIDYVQKQYREGCAQIYTF-QKF-KTR-SALRDVGKFNHLEGAEDLKG-LP 415
 QY 171 KGNLSLIDIDFNVHSGENQ-IOIVNFQKIYTVSDAYKNGEDVQDTYVIEDLKQKGIS 229
 DB 416 KDKSL 421
 QY 230 AERPLV 235

RESULT 14
 ID SLAP_BACST STANDARD: PRT: 1228 AA.
 AC P35825;

DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE S-LAYER PROTEIN PRECURSOR (SURFACE LAYER PROTEIN).
 GN SBA.
 US BACILLUS STEAROTHERMOPHILUS.
 CC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 CC BACILLUS.
 CC [1]
 RN RP
 RP SEQUENCE FROM N.A.

RC STRAIN-PV72;
 RX MEDLINE: 94320770.
 RA KPM R., SLEYTR U.B., LUBITZ W.;
 RT "Sequence analysis of the sba gene encoding the 130-kDa
 RT surface-layer protein of Bacillus stearothermophilus strain PV72."
 RL GENE 145:115-120(1984)
 CC -1- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
 CC OF PROTEINUS WHICH COAT THE SURFACE OF BACTERIA.
 CC -1- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIA IS COVERED BY A
 CC S-LAYER WITH HEXAGONAL SYMMETRY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X71092: G312730: -
 DR PIR: S34365; S34365.
 DR HSSP: P02766; IETB.
 KW SIGNAL: CELL WALL; S-LAYER.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 1 1228 S-LAYER PROTEIN.
 SQ SEQUENCE 1228 AA: 131076 MW: 0D0FB37F CRC32;

Query Match 3.3%; Score 110; DB 1; Length 1228;
 Best Local Similarity 24.7%; Pred. No. 1.97e-01;
 Matches 41; Conservative 43; Mismatches 72; Indels 10; Gaps 10;

DB 200 YDIIV-AMKAR-EV-QDAVKAAGNDKAKAAVDOINQYLPKVTDKFTLT-EVAKKALDA 255
 QY 200 YTVSDAVKNGEDVQDTYVIEDLKQKGISARPLVYISSVAYGRQYKLETTSKS-DE 256
 DB 256 DEALTPEVES-AINONKAVELTAVPVGTLKQLSLAANEDVYNNVTYIYVDG 313
 QY 259 VEAFAELIKGVKAPQTEWKI-LDNTKAVAILGGDPSGARGVTVGKVDVIEDLIQEG 317

DB 314 NIPRALTAADVSLSTDKITITVDASTPENNTEYKVVKGIDKNG 359
 QY 318 S-RETAHPGLPISTTSFLRDNVAVATFONSTDY-VEIKVTAVRNG 361

RESULT 15
 ID CASK_UNCUN STANDARD: PRT: 146 AA.
 AC Q29150;

DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE KAPPA CASEIN (FRAGMENT).
 GN CSN3 OR CSNK.
 OS UNCIA UNCIA (SNOW LEOPARD) (PANTHERA UNCIA).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA.
 OC CARNIVORA; FISSIPEDIA; FELIDAE; UNCIA.
 RN [1]
 RP SEQUENCE FROM N.A.

RA GATEY J., HAYASHI C., CROBIN M., ARCTANDER P.;
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: KAPPA CASEIN STABILIZES MICELLE FORMATION, PREVENTING
 CC CASEIN PRECIPITATION IN MILK.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
 CC -1- SIMILARITY: BELONGS TO THE KAPPA-CASEIN FAMILY.

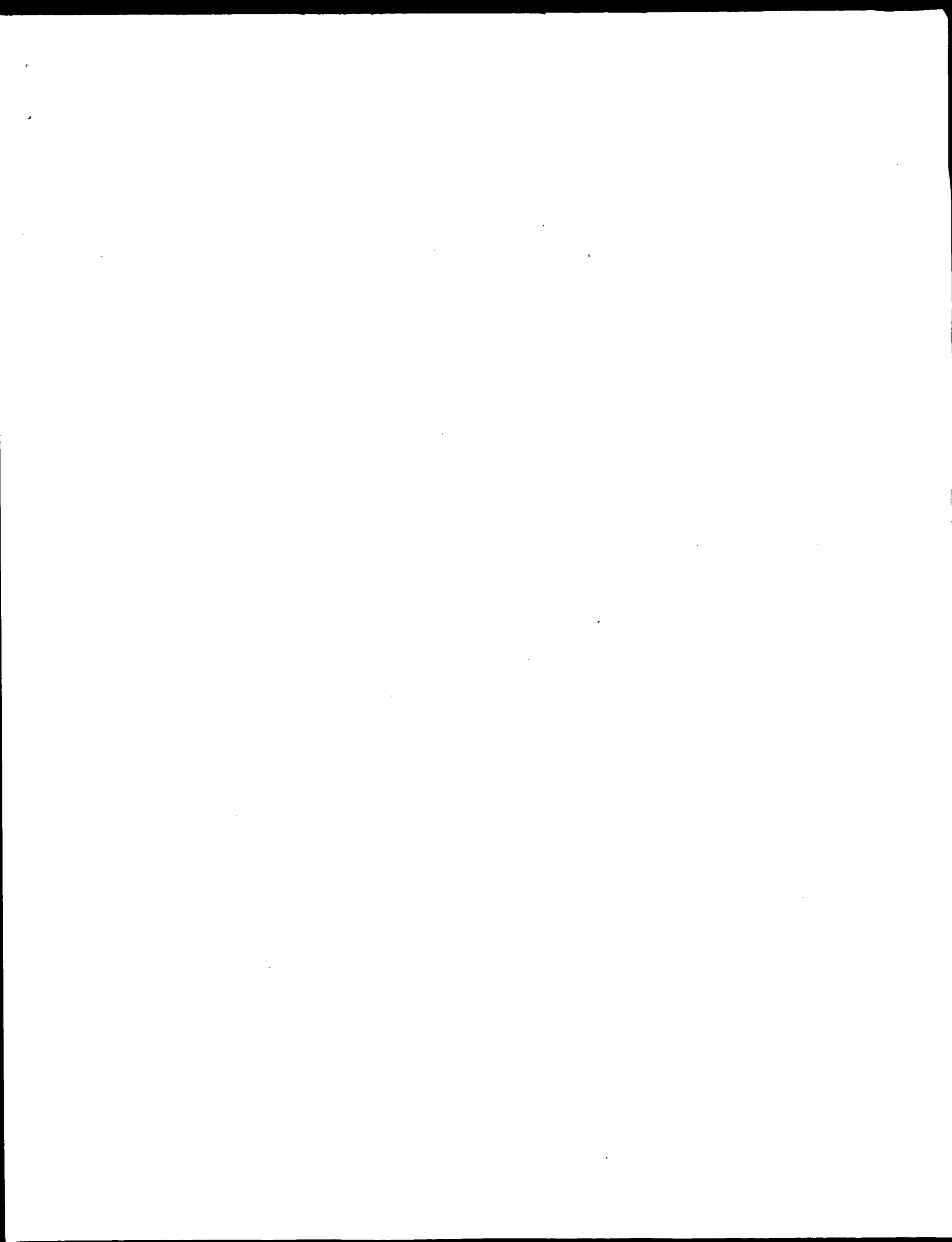
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U53894: G1293141: -
 DR PRAM: PR00997: casein_kappa.1
 KW MILK; PHOSPHORYLATION; GLYCOPROTEIN.
 FT NON-TER
 SQ SEQUENCE 146 AA: 16337 MW: 8321CFD3 CRC32;

Query Match 3.2%; Score 105; DB 1; Length 146;
 Best Local Similarity 25.0%; Pred. No. 8.42e-01;
 Matches 16; Conservative 19; Mismatches 29; Indels 0; Gaps 0;

DB 76 HUPASEIVIPKKIDKIDKIDKIDKIDKIDKIDKIDKIDKIDKIDKIDKIDKIDKID 135
 QY 38 QLPDEVVIERKKRSLSTSTSDISVATNDSRLYGCALLVDEITLNNPILAVDAPM 97
 DB 136 TTVV 139
 QY 98 TYSI 101

Search completed: Mon Aug 30 16:05:06 1999
 Job time : 28 secs.



 WISE (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_dp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 30 16:05:23 1999; MasPar time 29.10 Seconds

Tabular output not generated. 883.456 Million cell updates/sec

Title: >US-09-120-04-3

Description: (1-471) from US09120044.pep

Perfect Score: 3299

Sequence: 1 MANKAVNDITLAMYDKKKL.....TISWGTTLTPQVEDKVEN 471

Scoring table:

PAM 150

Gap 11

Searched: 179066 segs, 54579741 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

sptrembl9

1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
 5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
 9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
 13:sp-vertebrate 14:sp-virus

Statistics: Mean 49.610; Variance 95.041; scale 0.522

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1808	54.8	497	2	055996	SULIYSIN.	0.00e+00
2	1803	54.7	497	2	085102	HEMOLYSIN.	0.00e+00
3	1329	40.3	534	2	031241	POLYSIN.	4.77e-251
4	124	3.8	50	2	048772	LISTERIOLYSIN O (FRAGM	2.46e-03
5	118	3.6	50	2	048773	LISTERIOLYSIN O (FRAGM	1.69e-02
6	118	3.6	1361	2	030524	S-LAYER PROTEIN.	1.69e-02
7	118	3.6	1361	2	087083	S-LAYER PROTEIN.	1.69e-02
8	117	3.5	265	2	012044	CHROMOSOME XV READING	2.32e-02
9	114	3.5	450	2	068518	HYPOHETICAL 49.7 KD P	1.62e-01
10	111	3.4	831	3	013705	HYPOHETICAL 95.0 KD P	1.48e-01
11	109	3.3	867	10	081747	DNA GYRASE A SUBUNIT.	2.71e-01
12	110	3.3	867	10	081747	HYPOHETICAL 99.2 KD P	2.01e-01
13	106	3.2	187	1	058072	187AA LONG HYPOHETICA	6.62e-01
14	104	3.2	231	11	035813	KERATIN 14 (FRAGMENT).	1.19e+00
15	104	3.2	303	3	005541	CHROMOSOME IV COSMID 9	1.19e+00
16	106	3.2	501	1	028947	CONSERVED HYPOHETICAL	6.62e-01
17	107	3.2	710	2	085906	TRANSFERRIN PROTEIN B.	4.93e-01
18	106	3.1	1029	10	080955	TI6B24.10 PROTEIN.	6.62e-01
19	102	3.1	142	2	085634	L0038.	2.11e+00
20	102	3.1	277	2	051546	PHOSPHATE TRANSPORT AT	2.11e+00

21	101	3.1	362	2	050434	AMINOTRANSFERASE.	2.80e+00
22	102	3.1	446	2	034363	PUTATIVE L-AMINO ACID	2.11e+00
23	102	3.1	1329	5	076356	C45G7.6 PROTEIN.	2.11e+00
24	101	3.1	5027	11	063460	PROLINE-RICH PROTEIN (2.80e+00
25	100	3.0	124	5	017618	C04B4.6 PROTEIN.	3.72e+00
26	100	3.0	142	2	052137	ROPER.	3.72e+00
27	98	3.0	165	13	P87487	GROWTH HORMONE (FRAGME	6.50e+00
28	99	3.0	218	1	058691	HYPOHETICAL PROTEIN M	4.92e+00
29	99	3.0	259	1	027868	CELL DIVISION INHIBITO	4.92e+00
30	99	3.0	495	2	033330	HYPOHETICAL 53.3 KD P	4.92e+00
31	98	3.0	508	5	025917	MITOGEN-ACTIVATED PROT	6.50e+00
32	98	3.0	539	10	080914	MYOSIN HEAVY CHAIN-LIK	6.50e+00
33	100	3.0	640	5	009647	HYPOHETICAL 74.1 KD P	6.50e+00
34	98	3.0	688	5	045430	F32B6.2 PROTEIN.	6.50e+00
35	100	3.0	814	10	065715	HYPOHETICAL 91.9 KD P	6.50e+00
36	98	3.0	865	2	P73260	HYPOHETICAL 97.4 KD P	6.50e+00
37	100	3.0	1035	2	025887	CATION EFFLUX SYSTEM P	3.72e+00
38	99	3.0	1359	10	081726	HYPOHETICAL 155.3 KD	4.92e+00
39	99	3.0	1817	5	019931	COSMID F31D5.	4.92e+00
40	97	2.9	315	2	P70803	RESTRICTION ENDONUCLEA	8.56e+00
41	97	2.9	344	2	048027	OUTER MEMBRANE PROTEIN	8.56e+00
42	97	2.9	358	2	P74134	AT103.	1.12e+01
43	96	2.9	437	2	P74770	7-KETO-8-AMINOPELAGON	1.12e+01
44	96	2.9	1123	14	069089	UL37.	8.56e+00
45	97	2.9	2160	5	017709	H0514.2 PROTEIN.	8.56e+00

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	497 AA.
AC	055996	055996		
DT	01-NOV-1996	(TREMBLREL. 01, CREATED)		
DT	01-JAN-1999	(TREMBLREL. 09, LAST SEQUENCE UPDATE)		
DT	01-JAN-1999	(TREMBLREL. 09, LAST ANNOTATION UPDATE)		
DE	SULIYSIN.			
GN	SLY.			
OS	STREPTOCOCCUS SUTS.			
OC	BACTERIA, FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;			
OC	STREPTOCOCCUS.			
RC	SEQUENCE OF 1-71 FROM N.A.			
RC	STRAIN-P1/7;			
RA	SEGERS R.P.A.M., KENTER T., DE HAAN W.A.M., JACOBS A.A.C.;			
RT	"Characterization of the gene encoding suliyisin from Streptococcus			
RT	suts, and expression in field strains";			
RT	SUBMITTED (OCT-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; Z36907; E1334304; -			
SQ	SEQUENCE 497 AA; 54850 MW; D78ABEF CRC32;			

Query Match	54.8%; Score 1808; DB 2; Length 497;
Best Local Similarity	50.4%; Pred. No. 0.00e+00;
Matches	236; Conservative 113; Mismatches 117; Indels 2; Gaps 2;
Db	29 SKODINYOFSILTEPEQELTNEGEYIDNPATGMLNGRFVYLREKKNITNSADIA 88
Qy	2 ANKAVNDITLAMYDKKKLTHGSELENRPRIKGNLPDEFVIERKKSLSSTNSDIS 61
Db	89 VIDAKAANIYPGALLRADONLDNNPTLISARGDLTSLNPLGANGDSHTVANSPTS 148
Qy	62 VTATNDSRLYPGALLVYDELTENNPTLLAVDRAPMTYSIDLPGLASDSDFLEDPDSNS 121
Db	149 TVRGVNNLLSKRNNTYAGEGNTQALQYDETMAYMSQLTKFGTSFKIAPVDINF 208
Qy	122 SVRGAVNDLAKMHQDY-GOVNVPARMQYEKLTFAHSMELKVFSGDFKGTGNSIDIF 180
Db	209 DAVNSGEKOVIVNEKQIYYTVSVDEPESPKLAEETTEDLKRNQITVEPPVYSSV 268
Qy	181 NSVHSGEKQIIVNEKQIYYTVSDAVKNGVDVFDVTVEDLKORQISAEPLVYISSV 240
Db	269 SYGSMFKLETSSRSRTQVQAAFKAGVDSIGNAEYQDILKNTSSAIFGSDGASAA 328
Qy	241 AYGRQVYIKLETTSKSDSEVAEFAELIKGVKAPQTEWKRQILNTEVKVNILOGDSSGA 300

Query Match	54.7%;	Score 1803;	DB 2;	Length 497;
Best Local Similarity	50.4%;	Pred. No. 0.00e+00;		
Matches 236;	Conservative 112;	Mismatches 118;	Indels 2;	Gaps 2;

RESULT	3		
ID	031241	PRELIMINARY;	PRT; 534 AA.
AC	031241;		

Query Match	40.3%;	Score 1329;	DB 2;	Length 534;
Best Local Similarity	40.7%;	Pred. No. 4,77e-251;		
Matches	192;	Conservative	125;	Mismatches 147;
			Indels	8;
			Gaps	

RESULT	4	
ID	Q48772	PRELIMINARY; PRT; 50 AA

DT 01-NOV-1996 (TREMBLREL, 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL, 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL, 08, LAST ANNOTATION UPDATE)
 DE LISTERIOLYSIN O (FRAGMENT).
 GN HLX.
 OS LISTERIA MONOCYTOGENES.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE.
 OC LISTERIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-11984, TYPE 1;
 RX MEDLINE; 96118685
 RA RASMUSSEN O.F., SKOUBE P., DONS L., ROSSEN L., OLSEN J.E.;

01-NOV-1996 (TREMBLREL, 01, CREATED)
01-NOV-1996 (TREMBLREL, 01, LAST SEQUENCE UPDATE)
01-NOV-1996 (TREMBLREL, 01, LAST ANNOTATION UPDATE)

DE CHROMOSOME XV READING FRAME ORF YOR220W.
 OS CHROMOSOMES CEREBISIAE (BAKER'S YEAST).
 OC EUKARYOTA: FUNGI: ASCOMYCOTA: HEMIASCOCYCETES; SACCAROMYCETALES;
 OC SACCAROMYCETACEAE; SACCAROMYCETES.
 RN
 RP SEQUENCE FROM N.A.
 RA BOYER J., FAIRHEAD C., GAILLON L., GALLISSON F., MICHAUX G.,
 RA THIERRY A., DUJON B.;
 RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MUPS;
 RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FY1679;
 RA GALLISSON F., DUJON B.;
 RL SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: Z75128; E252085; -;
 DR EMBL: X92441; G1050772; -;
 SO SEQUENCE 265 AA; 29255 MW; 2D5DD78E CRC32;
 Query Match 3.5%; Score 117; DB 3; Length 265;
 Best Local Similarity 27.6%; Pred. No. 2.32e-02;
 Matches 29; Conservative 28; Mismatches 41; Indels 7; Gaps 7;
 Db 85 PMKLFVETSLNNQ-HPR-SRSTDVAISLDNNLALLEDHNRKPLSINTDPGVTGYDSS 142
 Oy 12 AMN-YDKKKLLTHOGESEIENFKEGNOPLDE-FVYIE-RKKRSLSTNSISVTAINDS 68
 Db 143 SLNKGSSLPDKSSLESPLMKLSTDSKPSYODPLPKSRSSS 187
 Oy 69 RLY-FGALLVVDLELLENNPTL-LAVDRAPMTYSIDLPGLASDSS 111
 RESULT 9
 ID 068518 PRELIMINARY; PRT; 450 AA.
 AC 068518;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 OS MYXOCOCCUS XANTHUS.
 OS MYXOCOCCUS XANTHUS.
 OC BACTERIA: PROTEOBACTERIA: DELTA SUBDIVISION: MYXOBACTERIA;
 OC MYXOCOCCALES; CYSTOBACTERINEAE; MYXOCOCCACEAE; MYXOCOCCUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DZFL;
 RA TRUDEAU K.G., WARD M.J., ZUSMAN D.R.;
 RL "Identification and characterization of Frlz, a novel response
 RL regulator necessary for swarming and fruiting-body formation in
 RL Myxococcus xanthus.";
 RL MOL. MICROBIOL. 20:645-655(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DZFL;
 RA WARD M.J., ZUSMAN D.R.;
 RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF049107; G2947295; -;
 KW HYPOTHETICAL PROTEIN.
 SO SEQUENCE 450 AA; 49686 MW; E3A6FC67 CRC32;
 Query Match 3.5%; Score 114; DB 2; Length 450;
 Best Local Similarity 26.8%; Pred. No. 5.91e-02;
 Matches 30; Conservative 28; Mismatches 47; Indels 7; Gaps 6;
 Db 9 PLVLYVSDMTFAKL-LFADPKRGVHEHYLLATLSGSEIVPPD-PPIPLPSG-RLV 65
 Oy 233 PLVLYIS-VAYGRQVYLKLETSKSDVEADAEALIKVKAHPQTEWKQI-IDNTEVKAV 290
 Db 66 HLPGRPLVGLHPETGELELVREKVGKTFVFNAGALLPQYTRFLPGEV 117

Oy 291 ILGGDPSSGARVTVGKVDVMDLIQEGSRETFADHPG--LPISVTSFLRDNV 340
 RESULT 10
 ID 013705 PRELIMINARY; PRT; 831 AA.
 AC 013705;
 DT 01-JAN-1999 (TREMBLREL. 09, CREATED)
 DT 01-JAN-1999 (TREMBLREL. 09, LAST SEQUENCE UPDATE)
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 95.0 KD PROTEIN C13F5.06C IN CHROMOSOME I.
 GN SPAC13F5.06C.
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
 OC EUKARYOTA: FUNGI: ASCOMYCOTA: ARCHIASCOMYCETES;
 OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
 OC SCHIZOSACCHAROMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA BROWN D., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WOOD V.;
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: SOME, TO YEAST SEC10.
 DR EMBL: Z69091; E341164; -;
 KW HYPOTHETICAL PROTEIN: TRANSMEMBRANE.
 FT TRANSMEM 62
 FT TRANSMEM 82
 FT TRANSMEM 539
 FT TRANSMEM 559
 FT TRANSMEM 602
 FT TRANSMEM 622
 FT TRANSMEM 631
 FT TRANSMEM 651
 FT TRANSMEM 723
 FT TRANSMEM 743
 FT TRANSMEM 769
 FT TRANSMEM 789
 FT TRANSMEM 799
 SO SEQUENCE 831 AA; 94966 MW; 6F1C32E5 CRC32;
 Query Match 3.4%; Score 111; DB 3; Length 831;
 Best Local Similarity 18.2%; Pred. No. 1.48e-01;
 Matches 37; Conservative 70; Mismatches 84; Indels 12; Gaps 12;
 Db 495 IOENLIQSGFTEDYITIEKNSHLSNEKV-YSPFGHAAALPARASILISSODLSVYS 553
 Oy 28 IENRFKEBNQLPDEFVIERKKRSLSTNSDISTATNDSRLTGLALLVDELENNP 87
 Db 554 SYVNLIDKLIRDY-V-FKELSISQSYISHDKSNLDLHYLVDIRCKTIMGYFSAYIM 611
 Oy 88 TLLAV-DR-APMTYSIDLPGLASDSDFLQVEDPSNS-SVRGAVNDLLAKWQDYGOWNV 144
 Db 612 SYIPTGYTASRRETVNI-LSSISVIECAVNDVFATYHAQDHLEILLSYROI 670
 Oy 145 PARMOYEKTIITASHM-EQLVKFGSDFEKTGNL-DIDENSVAHS-GEK-QIOIVNFKQIYI 200
 Db 671 AMTEQIDSTELROSMTRNVON 693
 Oy 201 TVSVDAVKNKPGVFDQ-TVTVED 222
 RESULT 11
 ID 067108 PRELIMINARY; PRT; 744 AA.
 AC 067108;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE DNA GYRASE A SUBUNIT.
 GN GYRA.
 OS AQUIFEX AEOLICUS.
 OC BACTERIA: AQUIFICALES; AQUIFICACEAE; AQUIFEX.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE: 98196666.
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBEER K., SNEAD M.A., KELLER M., AUYAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
 RL "The complete genome of the hyperthermophilic bacterium Aquifex
 RL aeolicus.";
 RL NATURE 392:353-358(1998).

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RA DICKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBERG R., SNEAD M.A., KELLER M., ADUAY M., HOBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RA SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AE000716; G2983485; -
SQ SEQUENCE 744 AA; 84080 MW; 93795328 CRC32;

Query Match
Best Local Similarity 27.28; Score 109; DB 2; Length 744;
Matches 22; Conservative 23; Mismatches 31; Indels 5; Gaps 4;

Db 453 KKLIVASEGRT-KVFETEEELVKYG--D-KRRTFGVGVKEGSIYAVLQDSIIP 508
18 KLLTHGSEIENRFTIEGNDLDEFVIRKRKSLSTNTSDISVATYNDRLYPALLV 77
QY 509 VEELPLEKAP-VVNILRVPT 528
78 VDETLLENPTLLAVDRAPMT 98

RESULT 12
ID 081747; PRELIMINARY; PRT; 867 AA.
AC 081747;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 99.2 KD PROTEIN.
GN P16G20 210
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPARALE; BRASSICACEAE; ARABIDOPSIS.
RN [1]
RP SEQUENCE FROM N.A.
RA WATSON M.D., GIBBONS T., BARTLEY J., BANCROFT I., MEMES H.W.,
RA WAYER K.F.X., SCHUELLER C., BEVAN M.;
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RA EU ARABIDOPSIS SEQUENCING PROJECT;
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AL031326; E1316763; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 867 AA; 99176 MW; 4847C6E3 CRC32;

Query Match
Best Local Similarity 24.88; Score 110; DB 10; Length 867;
Matches 28; Conservative 26; Mismatches 51; Indels 9; Gaps 8;

Db 276 DEY-ETIRGRITTFERIESSIA-L-VFSDKYPESSKCLD-ELVEIKKMETSIYVF 331
136 QGYGVNNPAMQYKRTAHSMEDLKVFGSDFEKGTGSLDIDNSVHSGEKQIQIYNF 195
QY 332 -PIFYVKAESYKNOTGFRNVLLKTEEDVRRKVDNRISLLETEDMIWGRQ 384
196 KOIYTVSDAVKNGDVGFQDTV--TVEDLKOR-GISAEPLVYISSVAYG-RQ 245

RESULT 13
ID 058072; PRELIMINARY; PRT; 187 AA.
AC 058072;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE 18734 LONG HYPOTHETICAL PROTEIN.
GN PH0334.
OS PYROCOCUS HORIKOSHII.
OC ARCHAEA; EURYARCHAEOTA; THERMOCCOCALLES; THERMOCCOCACEAE; PYROCOCUS.
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-OT3;
RX MEDLINE: 98344137.
RA KAMARAVASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOVAMA A., NAGAI Y.,
RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMITA M., OHYUKU Y.,
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,
RA AKOI K., NAKAMURA Y., KOB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,
RA KIKUCHI H.;
RT "Complete Sequence and Gene Organization of the Genome of a
RT Hyper-Thermophilic Archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA RES. 5:55-76(1998).
DR EMBL: AP000002; D1030351; -
SQ SEQUENCE 187 AA; 21959 MW; 8EC990E3 CRC32;

Query Match
Best Local Similarity 34.08; Score 106; DB 1; Length 187;
Matches 18; Conservative 11; Mismatches 18; Indels 6; Gaps 4;

Db 123 IEYERQ-YGTLKVIKK--VNGTEALITEKTIIPVKMKV-ENPEIKTRV 170
240 VAYGRVYTKLETTSKSDVEAAFE-ALIKGVKVPOTETWKQIINDTEKRAVI 291
QY

RESULT 14
ID 035813; PRELIMINARY; PRT; 231 AA.
AC 035813;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE KERATIN 14 (FRAGMENT).
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCURCORNATHI; MORIDAE; MORINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-ESOPHAGUS;
RX MEDLINE: 97107478.
RA WANG D.-Y., XIANG Y.-Y., TANAKA M., SHEN Q., SUGIMURA H.;
RT "Identification of cyclokeratin subspecies altered in rat experimental
RT esophageal tumors by subtractive cloning."
RL CANCER LETT. 108:119-127(1996).
DR EMBL: D63774; D1023234; -
DR PFM: PF00038; filament; 1.
KW KERATIN.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 231 AA; 25684 MW; C908F8B0 CRC32;

Query Match
Best Local Similarity 22.58; Score 104; DB 11; Length 231;
Matches 20; Conservative 30; Mismatches 34; Indels 5; Gaps 5;

Db 102 VDNANVLQIDNRLA-ANDLRTRYNETSLRQVSDINNCVDEL-TMSRADLEMQ 159
83 LNNPTLLAVDRAPMTYSIDPLGLASSDSL-QVEDPSNSVAGAVNDLAKRHQDYG-Q 140
QY 160 IESITELALVKNHEEMALRVLRGSD 188
141 VNNVPARMQY-EKTAHSMEDLKVFGSD 168

RESULT 15
ID 005541; PRELIMINARY; PRT; 303 AA.
AC 005541;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE CHROMOSOME IV COSMID 9819.
GN D9819.6.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
RN [1]

```

```

SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA FULTON L.;
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA JONATHAN M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,
RL FAVELLO A., FULTON L., GATUNG S., GRECO T., KIRSTEN J., KUCABA T.,
RN HARTSMORTH K., HARKINS J., HILLER L., JIER M., JOHNSON D.,
RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,
RL MITLER N., NHAN M., PAULEY A., PELUSO D., RIEKEN L., RILES L.,
RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,
RL WILSON R., WATERSTON R.;
RN SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RP [3]
SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA WATERSTON R.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA JIA Y., CHERRY J.M.;
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; 051031; G1332636; -
SQ SEQUENCE 303 AA; 33964 MW; D494B483 CRC32;

```

```

Query Match 3 28; Score 104; DB 3; Length 303;
Best Local Similarity 29.98; Pred. No. 1.19e+00;
Matches 26; Conservative 26; Mismatches 27; Indels 8; Gaps 5;
DB 84 YGFELQGLPSKNNMNGNSNSNTNKSMEPELGHRAOKFILLNVPYHSHKNFDDFKILQS 143
QY 99 YSIDPLGLASDSEFLVEDPSNSVGVAVNDLAKWHQDYGVNNVPARMOYE--KI--T 154
DB 144 AHTEELIVTGEYIGDDIASGTSNILE 170
QY 155 AHSMEQLKVK--F-GSDPEK-TGNSLD 177

```

Search completed: Mon Aug 30 16:06:48 1999
 Job time : 85 secs.